(54) Title: REDUCING THE IMMUNOGENICITY OF FUSION PROTEINS

(57) Abstract: Disclosed are compositions and methods for producing fusion proteins with reduced immunogenicity. Fusion pro-
teins of the invention include a junction region having an amino acid change that reduces the ability of a junctional epitope to bind
to MHC Class II, thereby reducing its interaction with a T-cell receptor. Methods of the invention involve analyzing, changing, or
modifying one or more amino acids in the junction region of a fusion protein in order to identify a T-cell epitope and reduce its
ability to interact with a T-cell receptor. Compositions and methods of the invention are useful in therapy.

REDUCING THE IMMUNOGENICITY OF FUSION PROTEINS

5   **Related Applications**

[0001] This application claims priority to and the benefit of U.S. provisional patent application serial number 60/280,625, filed March 30, 2001, the entire disclosure of which is incorporated herein by reference.

Field of the Invention

10   [0002] The present invention relates generally to methods and compositions for making and using modified fusion proteins with reduced or no immunogenicity as therapeutic agents. More specifically, the invention relates to fusion proteins, made less immunogenic by identifying candidate T-cell epitopes and modifying the amino acid sequence to eliminate such epitopes.

15   **Background of the Invention**

[0003] Many therapeutic proteins are normal human proteins. For example, interleukin-2, erythropoietin, and growth hormone are all human proteins that are given to humans who already usually make endogenous levels of these proteins. In general, immune responses against completely normal human proteins are rare when

20   these proteins are used as therapeutics.

[0004] Recently it has become apparent that many fusion proteins with artificial activities are useful as therapeutic proteins. For example, Enbrel is a fusion of the extracellular domain of a TNF receptor with an IgG1 Fc region. Enbrel is used to treat rheumatoid arthritis, and is thought to function by titrating TNF and preventing

25   TNF action. However, a significant incidence of anti-Enbrel antibodies have been noted in patients treated with Enbrel.

[0005] Another example of a therapeutically useful class of fusion proteins is the immunocytokines. These proteins include an antibody moiety and a cytokine moiety, and are useful for targeting cytokines to diseased cells, such as cancer cells.

30   However, the therapeutic use of many of these fusion proteins is curtailed due to their immunogenicity in mammals, especially humans.

[0006] Therefore, there is a need to generate fusion proteins with reduced immunogenicity in order to use these proteins in therapy.

Summary of the Invention

[0007] The present invention features methods and compositions useful for producing fusion proteins with reduced immunogenicity for use in therapy. For example, the invention features immunocytokines, immunofusins, immunoligands, other antibody and Fc fusion proteins, cytokine-cytokine fusion proteins, and albumin fusion proteins with decreased immunogenicity.

[0008] The invention relates in part to the insight that fusion proteins contain sequences that are "non-self." For example, even in a fusion between two human proteins, the region surrounding the fusion junction comprises a peptide sequence that is not normally present in the human body. For example, a protein drug such as Enbrel is derived from two normal human proteins: TNF receptor and IgG1. However, the junction between TNF receptor and IgG1 is a peptide sequence that is not normally found in the human body.

[0009] Preferred methods of the invention involve reducing the immunogenicity of a fusion protein by reducing the ability of a junctional epitope (junctional peptide) to interact with a T-cell receptor by reducing its ability to bind (its binding affinity) to MHC molecules. According to the invention, the junctional epitope or peptide is preferably "non-self." In general, proteins, including therapeutic proteins, are immunogenic, in part because proteins are endocytosed by antigen-presenting cells and proteolyzed, and the resulting peptides bind to molecules called major histocompatibility complex (MHC) that present the peptides to T cells. The antigenic peptide – MHC complex on the surface of an antigen presenting cell (APC) activates T-cells to proliferate, differentiate and release cytokines. In parallel, B-cell differentiation and antibody production is induced which may further limit the therapeutic protein's effectiveness due to clearance. Thus, the antigenic peptide, if derived from a therapeutic protein, is capable of inducing a series of undesired immune responses. The therapeutic protein's effectiveness is limited due to titration by antibodies, and the induction of T-cell and B-cell responses is often deleterious due to inflammatory and allergic reactions in the patient.

[0010] The invention provides (1) the identification of novel amino acid sequences in the region of the immunoglobulin – target protein junction with one or more candidate T-cell epitopes; and (2) the modification of these amino acid sequences to reduce or eliminate the presence of peptides, derived from the junction sequence, that function as T-cell epitopes.

2

[0011] The invention provides two general classes of compositions and methods relating to the reduction of immunogenicity. According to one embodiment of the invention, potential non-self T-cell epitopes are identified in sequences that span a fusion junction. For example, potential non-self T-cell epitopes are identified by

5   computational methods based on modeling peptide binding to MHC Class II molecules. Substitutions are then made such that the ability of peptides deriving from the junction region to bind to MHC Class II is reduced or eliminated. This process of identifying and modifying peptides which bind to MHC Class II is termed "de-immunization" and the resultant modified protein molecules are termed "de-

10   immunized."

[0012] According to another embodiment of the invention, one or more glycosylation sites is introduced at a fusion junction. An N-linked glycosylation site is preferably used, although an O-linked glycosylation site may also be used. According to a preferred embodiment, amino acids in a junction region surrounding a fusion junction

15   of wild-type sequence are mutated such that the last amino acid of the N-terminal fusion partner is mutated to an asparagine, and the first two amino acids of the second fusion partner are mutated to a glycine followed by a serine or a threonine.

[0013] According to the invention, removal of MHC Class II binding is preferred in situations where a protein is to be produced in bacteria or in an organism that does not

20   generate a mammalian glycosylation pattern, such as yeast or insect cells.

[0014] The introduction of glycosylation sites may be preferred when the protein is to be produced in a mammalian cell line or in a cell line that creates a glycosylation pattern that is innocuous to mammals.

[0015] In a preferred embodiment, a component of the fusion protein is a cytokine.

25   The term "cytokine" is used herein to describe naturally occurring or recombinant proteins, analogs thereof, and fragments thereof that elicit a specific response in a cell that has a receptor for that cytokine. Preferably, cytokines are proteins that may be produced and excreted by a cell. Preferably, cytokines include interleukins such as interleukin-2 (IL-2), IL-3, IL-4, IL-5, IL-6, IL-7, IL-10, IL-12, IL-13, IL-14, IL-15,

30   IL-16 and IL-18, hematopoietic factors such as granulocyte-macrophage colony stimulating factor (GM-CSF), G-CSF and erythropoietin, tumor necrosis factors (TNF) such as TNFα, lymphokines such as lymphotoxin, regulators of metabolic processes such as leptin, and interferons such as interferon α, interferon β, and

interferon γ and chemokines. Preferably, the antibody-cytokine fusion protein of the present invention displays a cytokine specific biological activity.

[0016] In another preferred embodiment, a component of the fusion protein is an anti-obesity cytokine. For example, a component is leptin, CNTF, or a portion of Acrp30.

[0017] In an alternative preferred embodiment, a component of the fusion protein is a hormone. For example, a component may be insulin, growth hormone, or glucagon-like peptide 1(GLP-1).

[0018] In yet another alternative embodiment, a component of the fusion protein is a ligand-binding protein with biological activity. In a preferred embodiment, an extracellular domain of TNF receptor is used.

[0019] According to one series of embodiments, a fusion protein of the invention comprises the N-terminus of a non-antibody moiety fused to the C-terminus of an antibody moiety. According to another series of embodiments, a fusion protein of the invention comprises the C-terminus of a non-antibody moiety fused to the N-terminus of an antibody moiety. According to the invention, an antibody moiety can be an intact immunoglobulin or a portion of an intact immunoglobulin. A portion of an immunoglobulin can include a variable region or a constant region or both. Preferred immunoglobulins include Fc regions or portions thereof. A preferred embodiment of the invention includes an IgG1 immunoglobulin isotype, or a portion thereof, modified to be less immunogenic and/or to have a longer serum half-life. For example, an IgG1 with modification of amino acid residues near the CH3 – cytokine junction is preferred. For certain applications, antibody moieties from IgG2 or IgG4 isotypes are preferred.

[0020] Immunocytokines are only one example of a tumor-targeted fusion protein therapy. Other tumor-toxic molecules can also be targeted to tumors by fusion to tumor-specific antibodies. In addition, antibody fusion proteins can attack other types of diseased cells, such as virus-infected cells. Another approach to engineering targeted fusion proteins has been use of Fc-X and X-Fc technology where X is a polypeptide. These technologies utilize the knowledge that production and collection of a target protein is improved if the polypeptide of interest is linked to the Fc portion of an immunoglobulin. For Fc-X fusion proteins, a signal peptide, followed by the Fc fragment of an immunoglobulin gene is the N-terminal fusion partner to the target protein. In some instances it is specifically advantageous to engineer a fusion protein

in the X-Fc orientation. With these constructs the target protein is the N-terminal
fusion protein and the Fc fragment follows. For some proteins this approach is useful,
as has been shown with lymphocyte cell surface glycoprotein (LHR) (US patent
5,428,130), and glucagon-like peptide (GLP-1).

5   [0021] Accordingly, methods and compositions of the invention provide forms of Fc-
X and X-Fc fusion proteins with reduced-immunogenicity. According to the
invention, the immunogenicity of a fusion protein can be assayed according to a
method known in the art or disclosed herein.

[0022] Methods and compositions of the invention also provide albumin fusion

10   proteins with reduced immunogenicity. Human serum albumin (HSA), due to its
remarkably long half-life, its wide in vivo distribution and its lack of enzymatic or
immunological functions, has been used as a carrier for therapeutic peptides/proteins
(Yeh et al, PNAS 89:1904-1908, 1992). A genetic fusion of a bioactive peptide to
HSA is useful for recovery of a secreted therapeutic HSA derivative. However,

15   according to the invention, albumin fusion proteins such as HSA-CD4 have a novel
junction which generally contains one or more T-cell epitopes capable of being
presented on MHC class II molecules. The invention provides less immunogenic
forms of albumin fusion proteins, and general methods for reducing the
immunogenicity of albumin fusion proteins. According to the invention, useful

20   albumin proteins include species, allelic, and mutant variants of albumin, including
fragments thereof. Preferred albumin proteins retain the structural and functional
properties of a wild-type albumin protein such as HSA.

[0023] In another aspect, the invention provides de-immunized antibody fusion
proteins with normal, mutant, or hybrid isotypes that comprise useful mutations.

25   These mutations may be near the junction or at positions distinct from the region of
the junction.

[0024] For example, the invention provides a de-immunized immunocytokine,
modified at the junction, with a point mutation at the junction between the IgG and
non-IgG moieties. The cytokine moiety includes any cytokine but preferably IL-2 or

30   IL-12. In one embodiment, the amino acid changes involve changing the C-terminal
lysine of the antibody moiety to a hydrophobic amino acid such as alanine or leucine.
A key advantage of combining such mutations with a de-immunizing modification of
the invention is that the mutations act together to increase serum half-life and to
decrease immunogenicity. The methods described herein for combining de-

immunization of a fusion junction with a serum-half-life altering mutation are useful
to improve significantly the clinical efficacy of these fusion proteins.

[0025] In another aspect, the invention provides immunocytokines comprising a
hybrid antibody moiety that includes domains from different Ig isotypes, preferably
from both IgG1 and IgG2 isotypes, and a de-immunizing modification at the fusion
junction. For example, the invention provides a de-immunized, junction-modified
immunocytokine using an IgG2 and an IgG2h hybrid (IgG2 modified in the hinge
region to IgG1). In a preferred embodiment, the hybrid fusion protein consists of a
de-immunized immunoglobulin moiety composed of an IgG (γ1:CH1-H)(γ 2: CH2-
CH3) and a cytokine moiety.

[0026] In another aspect, the invention provides novel nucleic acid sequences that
encode fusion proteins with reduced immunogenicity or facilitate the expression,
production, and secretion of fusion proteins with reduced immunogenicity. Such
nucleic acids are generated according to standard recombinant DNA techniques.

[0027] In a preferred embodiment, a nucleic acid molecule encodes an
immunocytokine fusion protein. A preferred immunocytokine includes a cytokine,
for example, Interleukin 2, and a tumor specific monoclonal antibody such as an
antibody to human epithelial cell adhesion molecule KSA (EP-CAM)(huKS).

[0028] In another preferred embodiment, nucleic acid molecules encode Fc fusion
proteins in various configurations. The nucleic acid molecule encodes serially in a 5'
to 3' direction, (i) a signal sequence, an immunoglobulin Fc region and a target
protein sequence or (ii) a signal sequence, a target protein, and an immunoglobulin Fc
region, or (iii) a signal sequence, a first target protein, an immunoglobulin Fc region,
and a second target protein. The resulting nucleic acid molecule thereby encodes an
Fc-X, X-Fc, or X-Fc-Y structure where X and Y are a target protein. In an alternative
embodiment, a nucleic acid encodes an Fc-X, X-Fc, or X-Fc-Y protein without a
signal sequence.

[0029] In another preferred embodiment, a nucleic acid of the invention encodes an Ig
fusion protein with mutant or hybrid isotypes. Specifically, the nucleic acid provides
antibody moieties with hybrid isotypes, or alternatively with altered hinge regions.
For example, the fusion protein consists of an IgG2, modified to contain fewer
disulfide bonds in the hinge region, or an IgG2 CH2 and CH3 region in which the

hinge region derives from another antibody, preferably a normal or mutant IgG1
hinge region.

[0030] A nucleic acid of the invention is preferably incorporated in operative
association into a replicable expression vector which is then introduced into a
mammalian host cell competent to produce the fusion protein. The resultant fusion
protein is produced efficiently and secreted from the mammalian host cell. The
secreted fusion protein is subsequently collected from the culture media without
lysing the mammalian host cell. The protein product is assayed for activity and/or
purified using common reagents as desired, and/or cleaved from the fusion partner, all
using conventional techniques.

[0031] Thus, the invention also provides methods for producing fusion proteins with
reduced immunogenicity.

[0032] Methods and compositions of the invention are also useful to provide
therapeutic treatment using a fusion protein that has been rendered less immunogenic.
An overall object of the invention is to provide processes that are both efficient and
inexpensive as well as proteins that are less immunogenic. Preferred therapeutic
compositions of the invention include a therapeutically effective amount of de-
immunized fusion protein. Preferably, the de-immunized fusion protein is
administered along with a pharmaceutically acceptable carrier.

[0033] The foregoing and other aspects, features and advantages of the present
invention will be made more apparent from the detailed description, drawings, and
claims that follow.

<div align="center">**Detailed Description of the Invention**</div>

[0034] All proteins, including antibodies, that are administered to a patient for
therapeutic use have the potential to induce an immune response in the recipient host.
This immune response is mediated by T-lymphocytes (T-cells) which then trigger B-
lymphocytes (B-cells) to make antibodies. Antibody production against the
therapeutic agent is detrimental since it leads to more rapid elimination of the
therapeutic agent and may induce an allergic response.

[0035] The present invention provides methods of reducing the immunogenicity of
fusion proteins. According to one method of this invention, potential T-cell epitopes
are identified in the junction region of a fusion junction in a fusion protein. T-cell
epitopes are identified by a variety of computer and non-computer methods, including
prediction based on structure-based computer modeling or by synthesis of peptides

and testing for binding to specific MHC Class II molecules or in an immunogenicity
assay.

[0036] According to the invention, a fusion junction is defined as the position
between the last (C-terminal) amino acid of a first protein or peptide and the first (N-
5   terminal) amino acid of a second protein or peptide in a fusion protein. Accordingly,
a fusion junction includes any amino acids between the last amino acid of one protein
and the first amino acid of a second protein. In one embodiment, the fusion junction
includes a linker.

[0037] According to the invention, a junction region is the region of a fusion protein
10   surrounding or spanning the fusion junction between two proteins. A junction region
preferably includes between 1 and about 100 amino acids, more preferably between 1
and about 50 amino acids, or between 1 and about 25 amino acids, and even more
preferably between 1 and about 15 amino acids, or between 1 and 9 amino acids. In
one embodiment, a junction region comprises a spacer or linker peptide inserted at the
15   junction point between the two proteins. According to the invention, a junction
region including a spacer or linker peptide can also be de-immunized to minimize the
response of a patient to a fusion protein including the spacer or linker.

[0038] According to the invention, a junctional T-cell epitope is defined as a peptide
sequence capable of binding an MHC Class II containing at least one amino acid
20   derived from each of at least two different fusion partner proteins. For example, Paul
(*Fundamental Immunology*, Chapter 8, Table 8, p. 276 [2000] 4th ed.) illustrates
segments of 10 amino acids that can bind to an MHC Class II molecule. In a
junctional T-cell epitope, these 10 amino acid peptides are derived from different
fusion partners. According to the invention a potential or candidate T-cell epitope
25   that spans a fusion junction (a candidate junctional T-cell epitope) preferably includes
1 to 8 amino acids from either side of the junction, and more preferably 1 to 10 or 1 to
11 amino acids from either side of the junction. Candidate epitopes are preferably 9,
11, or 12 amino acids long. Accordingly, since a junctional T-cell epitope of the
invention includes at least one amino acid from each side of the junction, preferred
30   candidate T-cell epitopes are junctional epitopes that include 1-8 (or 1-10, or 11)
amino acids from one side of the junction and also include a complementary number
of amino acids from the other side of the junction to result in an epitope having 9-12
amino acids, and most preferably 9 amino acids.

[0039] According to the invention, anchor residues within a junctional T-cell epitope are then mutated to prevent binding to an MHC Class II molecule. In general, care is taken to not introduce additional potential T-cell epitopes, and to preserve the function of each fusion partner.

5 [0040] According to the invention, a fusion of wild-type sequences is a fusion in which the sequences at the N-terminal and C-terminal sides of the fusion junction are derived directly from naturally occurring sequences.

[0041] According to the invention, a de-immunized fusion junction is a junction sequence in which one or more substitution mutations have been introduced relative

10 to a junction of wild-type sequences. In a most preferred embodiment, deimmunization of a fusion junction does not involve introduction of a linker, such as a 'non-immunogenic' Gly-Ser linker, and the spatial relationship between the fusion partners is not altered in a de-immunized fusion protein. According to the invention, one or more amino acids can be substituted or changed in the junction region either

15 N-terminally to the fusion junction, C-terminally to the fusion junction, or both N-terminally and C-terminally to the fusion junction.

[0042] According to the invention, a potential T-cell epitope is a sequence that, when considered as an isolated peptide, is predicted to bind to an MHC Class II molecule or an equivalent in a non-human species. A potential T-cell epitope is defined without

20 consideration of other aspects of antigen processing, such as the efficiency of protein uptake into antigen-presenting cells, the efficiency of cleavage at sites in an intact protein to yield a peptide that can bind to MHC Class II, and so on. Thus, the set of T-cell epitopes that are actually presented on MHC Class II after administration of a protein to an animal is a subset of the potential T-cell epitopes.

25 [0043] According to the invention, a T-cell epitope is an epitope on a protein that interacts with an MHC class II molecule. Without wishing to be bound by theory, it is understood that a T-cell epitope is an amino acid sequence in a protein or a fusion protein, that failed to undergo the negative T-cell selection process during T-cell development and therefore will be expected to be presented by an MHC Class II

30 molecule and recognized by a T-cell receptor. In a preferred embodiment of the invention, the non-self T-cell epitopes are present in the junction region at the fusion junction of two proteins that form a fusion protein.

[0044] The invention provides non-computer methods for reducing or eliminating the number of T-cell epitopes in a fusion protein junction without requiring elaborate computer simulations or protein three-dimensional structures. In one embodiment, a method of the invention takes advantage of the fact that a core segment of nine amino

5   acids interacts with both the MHC class II molecule as well as the T-cell receptor during antigen presentation. The N-terminal most amino acid is called an "anchor" position residue that binds to a deep pocket within the MHC class II molecule. One of the following amino acids is typically present at the anchor position which is important for binding to an MHC class II molecule: Leucine, Valine, Isoleucine,

10   Methionine, Phenylalanine, Tyrosine and Tryptophan. According to the invention, an additional 2 to 3 amino acids adjacent to the core 9 amino acids also affect the interaction with MHC molecules. In addition, the C-terminal most amino acid in the first protein of the fusion protein can generally be substituted. This is useful especially when the N-terminal fusion partner or first protein is known to be active

15   when fused to the C-terminal fusion partner or second protein at the C-terminus of the first protein.

[0045] A general method of the invention includes mutating any Leucines, Valines, Isoleucines, Methionines, Phenylalanines, Tyrosines or Tryptophans that occur in the C-terminal most eight amino acids of an N-terminal fusion partner in a fusion protein.

20   In one embodiment, one or more of these amino acids in a candidate junctional T-cell epitope amino acids is preferentially mutated to a Threonine, an Alanine or a Proline. This retains some of the hydrophobic nature of the amino acid that is replaced. In further embodiments of the invention, one more more of the above-mentioned amino acids is deleted from a candidate or potential junctional T-cell epitope, or replaced

25   with an appropriate amino acid analog. According to the invention, if an amino acid is deleted to destroy a potential T-cell epitope, care is taken not to generate a new T-cell epitope that includes amino acids near the deletion.

[0046] According to the invention, it is often useful to construct a generalized expression plasmid construction intermediate comprising the coding sequence for an

30   N-terminal fusion partner containing a mutation of one or more hydrophobic residues in the last eight amino acids. Generally, such a plasmid has one or more convenient restriction enzyme sites at or near the DNA encoding the C-terminus of the N-terminal fusion partner.

[0047] The purpose of a plasmid construction intermediate is to construct expression plasmids encoding a fusion protein in which one or more N-terminal fusion partners has one or more substitutions of a Leucine, Valine, Isoleucine, Methionine, Phenylalanine, Tyrosine, or Tryptophan to another amino acid in the eight C-terminal

5   amino acids. The construction of such final expression plasmids may be accomplished by a variety of other methods well known in the art, such as generation of PCR fragments or synthetic nucleic acids, followed by ligation of the fragment into an appropriated vector or attachment with other sequences through well-known PCR techniques.

10   [0048] Specific preferred embodiments include Fc-X fusion plasmids, albumin-X fusion plasmids, scFv-X fusion plasmids, and Fab-X fusion plasmids. In the Fc(gamma)-X case, it is useful to introduce mutations into the coding sequence to bring about amino acid substitutions of the Leucine-Serine-Leucine-Serine segment near C-terminus the Fc region of an IgG1, IgG2, IgG3, or IgG4 molecule, as

15   diagrammed here for IgG1: Amino acid sequences of human Fc regions derived from IgG1, IgG2, IgG3 and IgG4 are depicted in SEQ ID NOs: 1, 2, 3 and 4 respectively.
        [0049] In one example, KSLSLSPGK (SEQ ID NO: 5) is changed to KSATATPGK (SEQ ID NO: 6). This mutation is designed to eliminate potential junctional T-cell epitopes and also remove a T-cell epitope in which the upstream Phenylalanine or

20   Tyrosine serves as a position 1 anchor residue.
        [0050] Alternatively, it is sometimes useful to combine mutations that remove candidate junctional T-cell epitopes with a mutation that extends the serum half-life. For example, by changing KSLSLSPGK (SEQ ID NO: 5) to KSATATPGA (SEQ ID NO: 7).

25   [0051] Other embodiments include substitutions in the LSLS segment to other amino acids such as Glycine or Proline.
        [0052] In the case of expression vectors used for making IgA fusion proteins, it is useful to delete some of the C-terminal amino acids, so that the cysteine near the C-terminus that is involved in oligomerization of IgA is deleted. For example, fifteen

30   amino acids can be deleted, such that the IgA heavy chain sequence ends with Proline-Threonine-Histidine before being fused to a second protein. In addition, it is useful to introduce the following changes near the C-terminus of CH3 domain of the IgA Fc region:
        QKTIDRLAGKPTH (SEQ ID NO: 8) changed to QKTADRTAGKPTH (SEQ ID NO: 9)

[0053] Additional de-immunized sequences in an IgA-X fusion protein are,

QKTPTRTAGKPTH (SEQ ID NO: 10)

QKTPTRPAGKPTH (SEQ ID NO: 11)

QKTATRPAGKPTH (SEQ ID NO: 12).

5

[0054] In the case of an albumin-X fusion, it is useful to introduce the following changes in an albumin-X expression plasmid such that the C-terminus of albumin is modified as follows:

KKLVAASQAALGL (SEQ ID NO: 13) changed to KKLVAASQAATTA (SEQ ID NO:

10   14).

[0055] Thus, the invention provides nucleic acid sequences and proteins that are useful in construction of less immunogenic fusion proteins. Specifically, the invention provides proteins with mutations of any Leucines, Valines, Isoleucines,

15   Methionines, Phenylalanines, Tyrosines, or Tryptophans in the last eight amino acids. The proteins are preferably human proteins with sequences that generally correspond to sequences found in the human body. The invention also provides nucleic acid sequences encoding such proteins. The nucleic acid sequences for this aspect of the invention may exist as plasmids, PCR-generated fragments, or nucleic acids produced

20   by chemical synthesis.

[0056] The invention also provides expression plasmids encoding a fusion protein in which one or more N-terminal fusion partners has one or more mutations of a Leucine, Valine, Isoleucine, Methionine, Phenylalanine, Tyrosine, or Tryptophan to another amino acid in the eight C-terminal amino acids.

25   [0057] For example, plasmids encoding an Fc-IL2 or whole-antibody-IL2 fusion protein in which the Fc region is mutated as described above are provided by the invention. In addition, fusions comprising an Fc region mutated as described above to normal or mutated forms of erythropoietin, such as the forms of erythropoietin described in WO01/36489, are provided by the invention.

30   [0058] The invention also provides a method for reducing immunogenicity of a fusion protein junction by introducing an N-linked or O-linked glycosylation site near, or preferably, at a fusion junction. For example, the amino acids Asparagine, Serine or Threonine, and a third residue are introduced as follows. Consider a sequence in

12

which X's represent amino acids of an N-terminal fusion partner, and Z's represent amino acids of a C-terminal fusion partner.

$X_1X_2X_3X_4X_5X_6Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9$

$X_1X_2X_3X_4X_5N\ G\ S\ Z_3Z_4Z_5Z_6Z_7Z_8Z_9$

5   [0059] According to this method, binding of a junction peptide is not necessarily blocked by introduction of the glycosylation site. However, any peptide that is bound in the MHC Class II groove and has the glycosylated asparagine C-terminal to the N-terminal-most anchor residue will not function as a T-cell epitope. The presence of the large glycosylation moiety will sterically hinder recognition of the MHC Class

10   II/peptide complex. A preferred glycosylation site includes the sequence Asn-X-Ser or Asn-X-Thr wherein X is preferably Gly, but can be any amino acid.
[0060] Furthermore, the introduction of mutations introducing Glycine and Serine residues does not create new T-cell epitopes. Neither Glycine nor Serine can act as an anchor residue. During antigen processing, a fusion protein, in principle, is cleaved

15   between the glycosylated Asparagine and the Glycine or between the Glycine and the Serine. In either case, the resulting peptides have the mutant Glycine and/or Serine residues N-terminal to an anchor residue, and thus the mutant Glycine and/or Serine residues are not recognized by a T cell receptor, since residues N-terminal to an anchor residue are outside the region recognized by the TCR.

20   [0061] In a variation of this method, a fusion junction region already contains a Serine or Threonine preceded by an amino acid residues such as Glycine, Serine, Alanine, etc. The second method is preferably used when a junction region is flexible and displaced from the hydrophobic core of each fusion partner, so that the novel N-linked glycosylation does not interfere with the folding or function of either fusion

25   partner.
[0062] It is a straightforward matter for those skilled in the art of protein engineering to determine when introduction of a glycosylation site is feasible. For example, the three-dimensional structure of each fusion partner, or close homologs of the fusion partners, may be known. It is often the case that a few amino acids at the N-terminus

30   or C-terminus of a protein are not resolved in an X-ray structure, or exhibit many possible conformations in an NMR structure. In cases where three or more amino acids are disordered on either side of a glycosylation site, there is some confidence that the resulting fusion protein will fold correctly and both partners will be active.

Some routine experimentation is necessary to determine whether a given fusion protein construct will be functional.

[0063] In preferred embodiments of the invention, both the N-terminal and the C-terminal partner of the fusion protein are human proteins. Potential T-cell epitopes in such fusion proteins are created from the final 8 amino acids of the N-terminal partner (first protein) combined with the first 8 amino acids of the C-terminal partner (second protein). This provides a series of 8 hybrid 9-mers created from the first and second proteins. Any aliphatic or aromatic residue (Leucine, Valine, Isoleucine, Methionine, Phenylalanine, Tryptophan or Tyrosine) in the last 8 amino acids of the first protein presents a high risk of creating an MHC binding peptide with the amino acid in the first position (anchor position) that binds the pocket of the MHC molecule. Therefore, substitution of any of the above-mentioned amino acids, with an amino acid that is not one of the above-mentioned amino acids, and preferably with Alanine, Proline, or Threonine, will remove a candidate T-cell epitope.

[0064] For example, in the case of an Fc fusion protein containing the sequence:

HNHYTQKSLSLSPGKGGGGSGGGGSGGGGS (SEQ ID NO: 15),

the leucine residues create two potential epitopes. Therefore, the sequence can be de-immunized as;

HNHYTQKSATATPGKGGGGSGGGGSGGGGS (SEQ ID NO: 16),

by changing L to A and S to T. These changes remove epitopes with Leucine as the first amino acid in the MHC binding pocket and Tyrosine as the first amino acid in the MHC binding pocket, respectively.

[0065] These substitutions for deimmunization work in humans for all Fc fusion proteins, both with and without linker sequences, preferably when 1) both proteins in the fusion protein are human proteins; 2) the MHC binding peptides in the natural sequences of both proteins are ignored; and 3) the 9-mers identical to the original sequences are also ignored.

[0066] Methods of the invention are generally applicable in all vertebrate organisms, preferably in mammals and most preferably in humans. The invention is illustrated further by the following non-limiting examples.

Examples

Example 1: Deduction of immunogenic reactive epitopes of huKS-IL2 immunocytokine.

[0067] HuKS-IL2 consists of humanized V_H and V_L regions combined with human H
and L chain constant regions. The H chain was fused at its carboxyl terminus to the
mature sequence of human IL-2 as described previously. This H chain is of the γ1
isotype and has high affinity for Fc receptors. Because of this high affinity HuKS-IL2
was cleared quickly from the circulation. Without wishing to be bound by theory, the
clearance of HuKS-IL2 presumably occurs via FcR-bearing cells in the liver (Kupffer
cells) and spleen (antigen presenting cells).

[0068] It was previously established that certain patients had made immune responses
to some portion of the huKS-IL2 molecule, however, the epitopes recognized by these
antibodies are not known. To deduce the reactive epitopes, relative reactivities of
patient sera with huKS-IL2 were compared to other related proteins:

(1) Hu14.18-IL2, a molecule with completely different humanized V regions
but exactly the same C regions and fusion junction with IL-2;

(2) VH1, a de-immunized form of huKS-IL2 with no T-cell epitopes in the VH
and VL regions, derived from mouse V regions with surface-exposed mouse B-cell
epitopes veneered to human residues.

(3) VH2, a de-immunized form of huKS-IL2 with one remaining T-cell
epitope in CDR3, derived from mouse V regions with surface-exposed mouse B-cell
epitopes veneered to human residues, in which the VH contains one T-cell epitope.

(4) 425-IL2 constructed with either KOL or EU Cγ1 regions (rather than KS)
(to compare allotypic reactivity);

(5) huKS-mIL2 – a molecule with the huKS V regions fused to mouse C
regions and mouse IL-2;

(6) human Fc-IL2;

(7) human Fc only;

(8) human IL-2 only.

[0069] Immunoglobulin fusion proteins and fragments were purified by protein A
Sepharose chromatography and were coated on 96-well plates in bicarbonate buffer
and then blocked with 1% goat serum containing 1% BSA. Dilutions of patient sera
were incubated and then unbound material was removed by three washes with PBS-

15

Tween. Bound human antibodies from the patient sera were detected with various
HRP-conjugated antibodies depending on the bound protein. Generally, goat anti-
human λ chain HRP conjugate was used because most of the plate-bound proteins
consisted of human Fc and human κ chains.

5   [0070] Certain patient sera showed a clear reactivity to huKS-IL2 that was not
detectable in pre-injection sera from the same patients. Preimmune antisera was used
to establish a baseline non-immunized control. Reactivity seen in patient sera can be
attributed to (1) anti-IL2 reactivity, (2) anti Fc (allotypic) reactivity, (3) reactivity to
the novel junction sequence or (4) anti-idiotypic reactivity with the KS idiotype, or a
10   combination of reactivities.

[0071] No patient serum reacted significantly with recombinant IL-2 or to
the Fc region (1 and 2 above). Some patients showed anti-idiotypic reactivity to the
KS V regions. All patient sera showed reactivity with Fc-IL2. Three of four patients
showed reactivity to Fc-IL2. The presence of reactivity against Fc-IL2 but not against
15   either Fc or IL2 suggests that the junction between Fc and IL2 was recognized by the
patients' anti-sera.

Example 2: Modification of amino acid residues at the junction of an antibody-
cytokine fusion protein to reduce immunogenicity by elimination of MHC Class II
20   binding motifs

[0072] Peptide threading analysis identified two overlapping peptide segments with
strong MHC binding potential at the junction between the Fc and IL2 portion of the
immunocytokine. The peptide threading and identification of potential T-cell
epitopes was performed as disclosed in Carr (WO00/34317). Amino acid changes
25   were introduced such that the existing potential MHC Class II binding epitopes were
eliminated, but new potential MHC Class II epitopes were not introduced.

[0073] Modification of a junction sequence LSLSPGK-AP (SEQ ID NO: 17) to
ATATPGA-AP (SEQ ID NO: 18)("LSLS to ATAT"), where the hyphen is the
immunocytokine huKS-IL2 junction, made junction-derived peptide sequences
30   incapable of binding to any human MHC Class II with an affinity high enough to
result in immunogenicity.

16

Example 3: Modification of amino acid residues at the junction of immunocytokine
fusion proteins to reduce immunogenicity

[0074] Modification of a junction sequence LSLSPGK-AP(SEQ ID NO: 17) to
LNLSPGA-AP (SEQ ID NO: 19)("LSLS to LNLS"), where the hyphen is the
immunocytokine huKS-IL2 junction, results in junction-derived peptide sequences
that are still capable of binding to certain MHC Class II molecules. However, when
the KS-IL2 protein is expressed in mammalian cells and secreted, the protein is N-
glycosylated near the junction because of the NXS/T sequence.

[0075] The resulting junction-derived peptides are not effective as T-cell epitopes,
because when the junction-derived peptides are presented to T cells by MHC Class II,
the large N-glycosylation moiety prevents specific docking between a T cell-receptor
and MHC Class II.

Example 4: Characterization of the immune reactivity of antigen presenting cells to
immunocytokine huKs-IL2 in comparison to a de-immunized huKS-IL2
immunocytokine.

[0076] Reduction of immunogenicity due to modification of the reactive epitope by
mutating LSLS to ATAT is directly tested as follows. Synthetic peptides mimicking
this sequence alter the immune response of a classic antigen presenting cell such as a
dendritic cell (DC). The following synthetic peptides
KSLSLSPGK-APTS (SEQ ID NO: 20)and
KSATATPGK-APTS (SEQ ID NO: 21),

where the hyphen is the KS-IL2 junction, are used to stimulate DC-mediated antigen
presentation to autologous T cells. The ability of those T cells to proliferate in
response to a subsequent challenge with the peptide antigen serves as a measure of
immunogenicity of that peptide.

[0077] Specifically, peripheral blood mononuclear cells (PBMC) are isolated from
leukopacks by standard density gradient techniques. Mononuclear cells are
resuspended in serum-free Aim V culture media and allowed to adhere. After 2 h at
37 °C nonadherent cells are removed. Adherent cells are cultured for 7 days in media
containing human GM-CSF (50 ng/ml) and IL-4 (20ng/ml) to derive immature
dendritic cells (DC). After 7 days, the cells are harvested and phenotypically

characterized by flow cytometry with appropriate FITC-labeled Abs for MHC class I, MHC class II, CD80 and CD40 to confirm the immature DC phenotype.

[0078] Non-adherent cells are cultured with IL2 and IL7 to obtain autologous effector cells (T-cells) to be used in subsequent functional studies. For functional studies, T-cells

5   are added to immature dendritic cells (10:1 ratio) and co-cultured with huKS, de-immunized huKS, peptide junction 13 mer (KSLSLSPGK-APTS) (SEQ ID NO: 20) and the modified, de-immunized 13 mer peptide (KSATATPGK-APTS) (SEQ ID NO: 21). Comparison of the proliferation index, as measured by tritiated thymidine incorporation after exposure to each of the immunocytokines or immunogenic and

10   modified de-immunized peptides demonstrates the degree of immunogenicity of each molecule. Namely, an increase in radioactive incorporation is roughly proportional to the ability of each peptide to be bind to a class II MHC molecule on DC and be presented to T cells.

15   **Example 5:** Deduction of immunogenic reactive epitopes found in albumin fusion proteins and modification of amino acid residues at a fusion junction to reduce immunogenicity.

[0079] Human serum albumin (HSA), due to its remarkably long half-life, its wide *in vivo* distribution and its lack of enzymatic or immunological functions, has been used

20   as a carrier for therapeutic peptides/proteins. A genetically engineered HSA-CD4 hybrid has been shown to block the entry of the human immunodeficiency virus into CD4+ cells while exhibiting antiviral *in vitro* properties similar to those of soluble CD4 (Yeh et al, PNAS 89:1904-1908, 1992). Thus, the genetic fusion of bioactive peptides to HSA is useful for designing and recovering secreted therapeutic HSA

25   derivatives. However, as with all fusion proteins, HSA-CD4 has a novel junction which can be immunogenic and contains T-cell epitopes capable of being presented on MHC class II molecules. Analysis of the junction between HSA and CD4 using the methods of Examples 1, 2, 3, and 4 identifies peptides with MHC binding potential. The potentially immunogenic sequences are modified to decrease or

30   eliminate potential T and B-cell epitopes in order to reduce immunogenicity. Similarly, a novel glycosylation site can be introduced into the junction region in order to reduce immunogenicity.

　　　　Albumin sequence　　　　　　　　　　CD4 sequence

　　　　TCFAEEGKKLVAASQAALGL – KKVVLGKKGDTVELTCTAS (SEQ ID NO:
22).

5   [0080] It is contemplated by the invention that the HSA-IFNalpha fusion protein junction
region contains three candidate T-cell epitopes,
KKLVAASQAALGL (SEQ ID NO: 13);
KLVAASQAALGLC (SEQ ID NO: 23); and
LGLCDLPQTHSLG (SEQ ID NO: 24).

10   [0081] The T-cell epitopes depicted in SEQ ID NOs: 13 and 23 overlap and can be
de-immunized by changing LV (in bold) to anything except F, I, L, M, V, W and Y.
Alternatively, the peptide threading score can be reduced significantly by changing
LG to TT. The T-cell epitope in SEQ ID NO: 24 can be de-immunized by changing
the second L (in bold) to an A.

15   [0082] Furthermore, it is contemplated that in the case of an HSA-X fusion, wherein
X can be any protein, deimmunization of the fusion junction is achieved by changing
the amino acid sequence AALGL(SEQ ID NO: 25) to TATTA (SEQ ID NO: 26).
CFAEEGKKLVAASQTATTA (SEQ ID NO: 27).

20   **Example 6: X-Fc fusion proteins and modification of amino acid residues at a fusion
junction to reduce immunogenicity.**
[0083] In some instances it is specifically advantageous to engineer a fusion protein
in the X-Fc orientation. With these constructs, a target protein is a N-terminal fusion
protein and a Fc fragment follows. For example, the glucagon-like peptide (GLP-1)

25   requires a free N-terminus for its activity, so a GLP-1-Fc fusion is useful.
[0084] A GLP-1-Fc fusion protein is constructed according to standard techniques
described in the art. This fusion protein has the C-terminus of GLP-1 joined to the
hinge of the γ1 heavy chain. The γ1 hinge sequence containing a Cys to Ser mutation
(residue 5) which eliminates the Cys residue that forms a disulphide bond with the

30   light chain in IgG1 (Lo et al., (1998) Protein Engineering 11:495-500) is used. The
non-mutant Fc sequence is
EPKSCDKTHTCPPCPAPELLG (SEQ ID NO: 28)

with the hinge region being underlined, followed by the start of the CH2 domain sequence.

[0085] The fusion junction between GLP-1 (7-37) and mutant Fc is:

HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG –

5 EPKS<u>S</u>DKTHTCPPCPAPELLG (SEQ ID NO: 29).

[0086] The fusion junction between GLP-1 (7-37) and normal Fc is:

SYLEGQAAKEFIAWLVKGRG – EPKSCDKTHTCPPCPAPELLG (SEQ ID NO: 30)

[0087] Three potential epitopes are identified by peptide threading at the GLP-1-Fc

10   fusion junction.

KEFIAWLVKGRGE (SEQ ID NO: 31)

EFIAWLVKGRGEP (SEQ ID NO: 32)

AWLVKGRGEPKSS (SEQ ID NO: 33).

15   [0088] Analysis of fusion junctions between GLP-1 (bold text) and Fc (plain text), performed as in Examples 1-3, identifies peptides with MHC binding potential. After identification of potential sites by peptide threading analysis, the potentially immunogenic sequences are modified by amino acid substitution to reduce or eliminate potential T and B-cell binding epitopes and decrease immunogenicity.

20   [0089] The above-mentioned potential T-cell epitopes depicted in SEQ ID NOs: 31, 32 and 33 are de-immunized by making single amino acid substitutions. For example, peptide shown in SEQ ID NO: 31 is de-immunized by changing the Lysine (shown in bold) to a Threonine and the Arginine(shown in bold) to a Threonine. The peptide shown in SEQ ID NO: 32 is de-immunized by replacing the Isoleucine (shown in

25   bold) with an Alanine or a Proline and the peptide in SEQ ID NO: 33 is de-immunized by replacing the Leucine with an Alanine or a Proline. The resulting de-immunized junction is:

HAEGTFTSDVSSYLEGQAAKEF<u>A</u>AW<u>A</u>V<u>T</u>G<u>T</u>G – EPKSSDKTHTCPPCPAPELLG

30   (SEQ ID NO: 34).

[0090] According to an exemplary method for introducing a glycosylation site at a fusion junction the following changes are introduced:

SYLEGQAAKEFIAWLVKGRN – GSKSSDKTHTCPPCPAPELLG (SEQ ID NO: 35).

Example 7: Deduction of immunogenic reactive epitopes of Enbrel, a TNFR-Fc
fusion protein and modification of amino acid residues at a fusion junction to reduce
immunogenicity.

[0091] ENBREL or etanercept, a X-Fc fusion protein approved by the FDA, is a
tumor necrosis factor (TNF) inhibitor used to treat rheumatoid arthritis. ENBREL is a
dimeric fusion protein consisting of an extracellular-ligand-binding domain of TNF
receptor linked to an Fc protein of human IgG1. TNFR-Fc competitively inhibits
binding of TNF to its receptor and renders the bound TNF biologically inactive,
resulting in significant reduction in inflammatory activity. As described above for
GLP-1 – Fc, TNFR-Fc has a novel junction which contains potential T-cell epitopes.

[0092] The junction between a direct fusion of a C-terminus portion of TNF-R **(bold
text)** to the N-terminus of the g1 hinge (plain text with the underline region
representing the hinge region) is

STSFLLPMGPSPPAEGSTGD – EPKSCDKTHTCPPCPAPELLG (SEQ ID NO: 36)

[0093] Analysis of a junction between TNF-R and Fc, performed as in Examples 1-4,
identifies peptides with MHC binding potential. After identification of potential sites
by peptide threading analysis, the potentially immunogenic sequences are modified by
amino acid substitution to reduce or eliminate potential T and B-cell binding epitopes
and decrease immunogenicity.

[0094] According to an exemplary method for introducing a glycosylation site at a
fusion junction the following changes are introduced:

STSFLLPMGPSPPAEGSTGN - GSKSCDKTHTCPPCPAPELLG (SEQ ID NO: 37).

Example 8: Deduction of immunogenic reactive epitopes for Fc-X-Y fusion proteins
such as Fc-IL12-IL2 and modification of amino acid residues at the fusion junction to
reduce immunogenicity.

[0095] Fusion proteins of a Fc-X-Y orientation such as Fc-IL12-IL2 have multiple
novel fusion junctions which are potentially immunogenic. For instance, Fc-IL12 has

21

a fusion junction similar to other Fc-X fusion proteins or immunocytokines (Example

1) but is novel due to the usage of the cytokine IL12. The fusion junction is analyzed

for immunogenic binding sites and modified accordingly. Secondly, there is an X-Y

fusion junction comparable to that described in Example 5, with two different

5   cytokines constituting a fusion protein. Peptide thread analysis is used for each of the

fusion junctions.

[0096] Analysis of the junctions:

(1) **MHEALHNHYTQKSLSLSPGK** - RNLPVATPDPGMFPCLHH SQ (SEQ ID NO:

    38)

10

between the C-terminus of Fc (bold text) and the N-terminus of IL12p35 (plain text),

and

(2) **RAQDRYYSSSWSEWASVPCS** - APTSSSTKKTQLQLEHLLLD (SEQ ID NO: 39)

15

between the C-terminus of IL12p40 (bold text) and the N-terminus of IL2 (plain text)

by peptide threading identifies peptides with MHC binding potential. The potentially

immunogenic sequences are modified to decrease or eliminate potential T-cell

epitopes.

20

[0097] For example, in sequence (1) above, the following changes are made:

MHEALHNHYTQKS<u>ATAT</u>PGK – RNLPVATPDPGMFPCLHHSQ (SEQ ID NO: 40).

25   [0098] These changes reduce or eliminate MHC Class II-binding potential of several

T-cell epitopes at a junction of Fc and the p35 subunit of IL12.

[0099] In another example, sequence (2) above is modified to introduce a

glycosylation site by introducing an Asparagine and Glycine at the first two positions

30   within IL-2. This strategy uses the naturally occurring Threonine at position 3 of

mature IL-2. In addition, it is important to not disrupt the formation of a disulfide

bond in the p40 moiety, so it is useful to separate the glycosylation site by at least one

or two amino acids from the Cysteine in p40.

RAQDRYYSSSWSEWASVPCS - <u>N</u>GTSSSTKKTQLQLEHLLLD (SEQ ID NO: 41).

[0100] In the case of the IL12p40-IL2 fusion, introduction of a glycosylation site as discussed above creates the following potential T-cell epitopes.

SEWASVPCSNGTS (SEQ ID NO: 42)

ASVPCSNGTSSST (SEQ ID NO: 43)

[0101] However, glycosylation of the T-cell epitope prevents MHC Class II binding thus resulting in reduced immunogenicity.

<u>Example 9</u>: <u>Deduction of immunogenic reactive epitopes in junction of an X-Fc-Y fusion protein and modification of amino acid residues at a fusion junction to reduce MHC Class II binding.</u>

[0102] Fusion proteins of the X-Fc-Y configuration, such as IL4-Fc-GMCSF, have multiple novel fusion junctions that contain potential T-cell epitopes. The IL4-Fc is a junction analogous to other X-Fc fusion proteins (Examples 6 and 7) but is novel due to the use of the cytokine IL4. For example, a form of Fc using a hinge region, CH2, and CH3 domain from human γ1 is used. As stated above, a γ1 hinge sequence in pdCs-huFcγ1 may contain a Cys to Ser mutation (underlined) that eliminates the Cys residue that forms a disulphide bond with a light chain in IgG1 (Lo et al., (1998) Protein Engineering 11:495-500), thereby creating a third potentially immunogenic fusion junction for analysis. The fusion junction is analyzed for potential T-cell epitopes and modified according to the methods of Examples 1-4.

[0103] There is an Fc-Y fusion junction comparable to that described in Example 1 for the immunocytokine huKS-IL2, with a different cytokine GMCSF constituting a fusion protein. This fusion junction is also analyzed for potential T-cell epitopes and modified according to the methods of Examples 1-4.

Specifically, analysis of the junctions

(1) **ENFLERLKTIMREKYSKCSS** – epks<u>c</u>dkthtcppcpapellg (SEQ ID NO: 44)

between the C-terminus of IL4 (bold text) and the N-terminus of Fc (plain text), and

(2) **MHEALHNHYTQKSLSLSPGK** – parspspstqpwehvnaiqe (SEQ ID NO: 45)

between the C-terminus of Fc (bold text) and the N-terminus of GMCSF (plain text) by peptide threading identifies peptides with MHC binding potential. The potential T-cell epitopes are modified to decrease or eliminate potential T epitopes in order to reduce immunogenicity.

[0104] A candidate T-cell epitope at the junction of IL4-Fc fusion protein is, EKYSKCSSEPKSC (SEQ ID NO: 46),

where changing E (in bold) to T reduces the peptide threading score or the MHC Class II binding potential significantly. The sequence of the modified IL4-Fc fusion is as follows:

ENFLERLKTIMREKYSKCSS – tpkscdkthtcppcpapellg (SEQ ID NO: 47).

[0105] The Fc-GMCSF fusion junction is de-immunized by changing the sequence LSLS to ATAT as shown below.

MHEALHNHYTQKSATAT PGK – parspspstqpwehvnaiqe (SEQ ID NO: 48).

Example 10: Modification of amino acid residues at a fusion junction of immunocytokines and immunofusins prepared with a hybrid isotype to remove T-cell epitopes.

[0106] It is often useful to construct an antibody or antibody-based fusion protein with a hybrid isotype, so that useful features of different isotypes may be combined into a single molecule. Fusion proteins with hybrid isotypes may be modified according to the invention to reduce immunogenicity.

[0107] An antibody fusion protein with the following components is constructed by standard recombinant DNA techniques: a light chain and a heavy chain, the V regions recognizing a tumor-specific antigen, the light chain being a typical light chain, and the heavy chain comprising CH1, CH2, and CH3 domains from IgG2 and a hinge region from IgG1, with a cytokine fused to the C-terminus of the heavy chain involving a fusion junction as described above.

[0108] This protein contains novel junctions between CH1g2 and hinge-g1, and hinge-g1 and CH2g2. The identification and modification of potential T-cell epitopes in these junctions is performed as follows. For immunocytokines and Fc-X fusion proteins prepared with either an IgG2 or an IgG2h isotype, these modifications are identical to those set forth in Examples 1, 2, 3, and 8 above. For X-Fc IgG2h immunofusins, the novel junction is also identical since the N-terminus of the Fc is located within the hinge · region of the IgG2h protein which has been modified to an IgG1 type. However, there are two novel fusion junctions in that the IgG1 hinge inserted into a IgG2 immunoglobulin creates two novel junctions between the IgG2 CH1 and IgG1 hinge and the IgG1 hinge and the IgG2 CH2.

IgG2 CH1 – IgG1 hinge –IgG2 CH2-IgG2 CH3 – target protein .

[0109] Thus, analysis of the junctions

qtytcnvdhkpsntkvdktv – epkscdkthtcppcp (SEQ ID NO: 49)

between the C-terminus of IgG2 CH1 (bold text) and the N-terminus of the IgG1 hinge (plain text), and

epkscdkthtcppcp – appvagpsvflfppkpkdtl (SEQ ID NO: 50)

between the C-terminus of the IgG1 hinge (bold text) and the N-terminus of IgG2 CH2 F (plain text) by peptide threading should identify peptides with MHC binding potential. The potentially immunogenic sequences are modified to decrease or eliminate potential T and B-cell epitopes in order to reduce immunogenicity.

[0110] Two potential T-cell epitopes in the IgG2CH1-IgG1 hinge fusion junction are,

TKVDKTVEPKSCD (SEQ ID NO: 51) and KTVEPKSCDKTHT (SEQ ID NO: 52).

[0111] The IgG2CH1-IgG1 hinge fusion junction is de-immunized by changing the V (in

bold) to an A, a T or a P. The sequence of the modified fusion junction is depicted in SEQ ID NO: 53.

qtytcnvdhkpsntkadkta – epks<u>c</u>dkthtcppcp (SEQ ID NO: 53).

[0112] As stated above, the γ1 hinge sequence in pdCs-huFcγ1 may contain a Cys to Ser mutation (underlined) that eliminates the Cys residue that forms a disulphide bond with the light chain in IgG1 (Lo et al., (1998) Protein Engineering 11:495-500), thereby creating two additional potentially immunogenic fusion junctions for analysis and modification:

(3) **qtytcnvdhkpsntkvdktv** – epks<u>S</u>dkthtcppcp (SEQ ID NO: 54)

(4)**epks<u>S</u>dkthtcppcp** – appvagpsvflfppkpkdtl (SEQ ID NO: 55).

Example 11: Generation of Fc-EPO fusion protein using hybrid isotype Fc components of IgG1 and IgG4.

[0113] To generate an Fc-erythropoietin fusion protein, the following expression plasmid was constructed using standard molecular biology techniques. An XmaI-XhoI DNA fragment containing a form of the human erythropoietin coding sequence with mutations resulting in the amino acid substitutions His32Gly, Cys33Pro, Trp88Cys, and Pro90Ala, as disclosed in WO01/36489, was used. The corresponding protein sequence is shown in SEQ ID NO: 56.

APPRLICDSRVLERYLLEAKEAENITTGCAEGPSLNENITVPDTKVNFYAWKRMEVGQQAVEV
WQGLALLSEAVLRGQALLVNSSQPCEGLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAAS
AAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGDR

[0114] This XmaI-XhoI DNA fragment was inserted into a plasmid vector that encodes a hinge region from IgG1 and a CH2 and CH3 region from IgG2, except that there were two sets of mutations that resulted in amino acid substitutions in the region of the CH3 C-terminus, such that the sequence at the junction of the CH3 C-terminus and the Epo N-terminus is as follows:

.... TQKSATATPGA-APPRLI ....(SEQ ID NO: 57)

[0115] The first set of mutations, which change the sequence KSLSLSPG (SEQ ID NO: 58) of the IgG2 CH3 region to KSATATPG (SEQ ID NO: 59), is disclosed in U.S. Patent Application Serial No. 60/280,625. The effect of the substitution of Leu-Ser-Leu-Ser (position 3 to position 6 of SEQ ID NO: 58) with Ala-Thr-Ala-Thr (position 3 to position 6 of SEQ ID NO: 59) is to remove potential human non-self T-cell epitopes that may arise because the junction between human Fc and human erythropoietin contains non-self peptide sequences. The second set consisting of the single amino acid substitution K to A at the C-terminal amino acid of the CH3 region, is disclosed in U.S. Patent Application Serial No. 09/780,668.

[0116] The resulting plasmid was transfected into NS/0 cells and the Fc-Epo fusion protein was expressed and purified according to the procedures known in the art. After purification based on binding to protein A, the huFcγ2h-huEpo protein containing the IgG2 CH3 and erythropoietin substitutions described above was characterized by size exclusion chromatography and found to consist of 97% monomer and 90% monomer in two independent preparations. The huFcγ2h-huEpo protein containing the IgG2 CH3 and erythropoietin substitutions described above was found to be about as active, on a molar basis, as human erythropoietin in a cell-based assay that measured the ability of an erythropoietin protein to stimulate TF-1 cell division. The assay was performed as described in WO01/36489.

[0117] In addition, fusions of non-mutant human erythropoietin to the C-terminus of an Fc region consisting of either IgG1(hinge-CH2-CH3), IgG2(hinge-CH2-CH3), or IgG1(hinge)-IgG2(CH2-CH3) were characterized. Expression plasmids comprising non-mutant human Fc sequences and non-mutant erythropoietin sequences were constructed analogously to the plasmid described above. NS/0 cells were transfected with the Fcγ1-Epo, Fcγ2-Epo, and Fcγ2h-Epo expression plasmids, and stable clones were isolated after screening an approximately equal number of clones for each plasmid. The best-producing clones yielded 50 μg/ml for Fcγ1-Epo, 20 μg/ml for Fcγ2-Epo, and 120 μg/ml for Fcγ2h-Epo.

27

[0118] The following example describes in detail a preferred method for identification of immunogenic sequence regions (T-cell epitopes) within the sequences of the fusion proteins as disclosed in this invention. However, it should be pointed out, that said molecules can be obtained by other known methods.

Example 12. Identification of T-cell epitopes by computational methods

[0119] According to the invention, epitopes in a junction region of a fusion protein can be modified using methods for introducing mutations into proteins to modulate their interaction with the immune system. According to the invention, known methods in the art that can be adapted according to the invention include those described in the prior art (WO 92/10755 and WO 96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699 755(Centro de Immunologia Moelcular), WO 98/52976 and WO 98/59244 (Biovation Ltd.) or related methods.

[0120] Advantageous mutant proteins, however, can be obtained if the identification of said epitopes is realized by the following new method which is described herewith in detail and applied to the junction region of fusion proteins according to the invention.

[0121] There are a number of factors that play important roles in determining the total structure of a protein, polypeptide or immunoglobulin. First, the peptide bond, i.e., that bond which joins the amino acids in the chain together, is a covalent bond. This bond is planar in structure, essentially a substituted amide. An "amide" is any of a group of organic compounds containing the grouping -CONH-.

[0122] The planar peptide bond linking Cα of adjacent amino acids may be represented as depicted below:
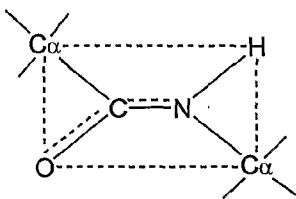
[0123] Because the O=C and the C-N atoms lie in a relatively rigid plane, free rotation does not occur about these axes. Hence, a plane schematically depicted by the interrupted line is sometimes referred to as an "amide" or "peptide plane" plane wherein lie the oxygen (O), carbon (C), nitrogen (N), and hydrogen (H) atoms of the peptide backbone. At opposite corners of this amide plane are located the Cα atoms. Since there is substantially no rotation about the O=C and C-N atoms in the peptide or amide plane, a polypeptide chain thus comprises a series of planar peptide linkages joining the Cα atoms.

[0124] A second factor that plays an important role in defining the total structure or conformation of a polypeptide or protein is the angle of rotation of each amide plane about the common Cα linkage. The terms "angle of rotation" and "torsion angle" are hereinafter regarded as equivalent terms. Assuming that the O, C, N, and H atoms remain in the amide plane (which is usually a valid assumption, although there may be some slight deviations from planarity of these atoms for some conformations), these angles of rotation define the N and R polypeptide's backbone conformation, i.e., the structure as it exists between adjacent residues. These two angles are known as ϕ and ψ. A set of the angles ϕ_i, ψ_i, where the subscript i represents a particular residue of a polypeptide chain, thus effectively defines the polypeptide secondary structure. The conventions used in defining the ϕ, ψ angles, i.e., the reference points at which the amide planes form a zero degree angle, and the definition of which angle is ϕ, and which angle is ψ, for a given

polypeptide, are defined in the literature. See, e.g., Ramachandran et al. *Adv. Prot. Chem.*
<u>23</u>:283-437 (1968), at pages 285-94, which pages are incorporated herein by reference.

[0125] The present method can be applied to any protein, and is based in part upon the
discovery that in humans the primary Pocket 1 anchor position of MHC Class II molecule
binding grooves has a well designed specificity for particular amino acid side chains. The
specificity of this pocket is determined by the identity of the amino acid at position 86 of
the beta chain of the MHC Class II molecule. This site is located at the bottom of Pocket
1 and determines the size of the side chain that can be accommodated by this pocket.
Marshall, K.W., *J. Immunol.*, <u>152</u>:4946-4956 (1994). If this residue is a glycine, then all
hydrophobic aliphatic and aromatic amino acids (hydrophobic aliphatics being: valine,
leucine, isoleucine, methionine and aromatics being: phenylalanine, tyrosine and
tryptophan) can be accommodated in the pocket, a preference being for the aromatic side
chains. If this pocket residue is a valine, then the side chain of this amino acid protrudes
into the pocket and restricts the size of peptide side chains that can be accommodated
such that only hydrophobic aliphatic side chains can be accommodated. Therefore, in an
amino acid residue sequence, wherever an amino acid with a hydrophobic aliphatic or
aromatic side chain is found, there is the potential for a MHC Class II restricted T-cell
epitope to be present. If the side-chain is hydrophobic aliphatic, however, it is
approximately twice as likely to be associated with a T-cell epitope than an aromatic side
chain (assuming an approximately even distribution of Pocket 1 types throughout the
global population).

[0126] A computational method embodying the present invention profiles the likelihood
of peptide regions to contain T-cell epitopes as follows: (1) The primary sequence of a
peptide segment of predetermined length is scanned, and all hydrophobic aliphatic and
aromatic side chains present are identified. (2) The hydrophobic aliphatic side chains are
assigned a value greater than that for the aromatic side chains; preferably about twice the
value assigned to the aromatic side chains, e.g., a value of 2 for a hydrophobic aliphatic
side chain and a value of 1 for an aromatic side chain. (3) The values determined to be
present are summed for each overlapping amino acid residue segment (window) of
predetermined uniform length within the peptide, and the total value for a particular

segment (window) is assigned to a single amino acid residue at an intermediate position of the segment (window), preferably to a residue at about the midpoint of the sampled segment (window). This procedure is repeated for each sampled overlapping amino acid residue segment (window). Thus, each amino acid residue of the peptide is assigned a value that relates to the likelihood of a T-cell epitope being present in that particular segment (window). (4) The values calculated and assigned as described in Step 3, above, can be plotted against the amino acid coordinates of the entire amino acid residue sequence being assessed. (5) All portions of the sequence which have a score of a predetermined value, e.g., a value of 1, are deemed likely to contain a T-cell epitope and can be modified, if desired.

[0127] This particular aspect of the present invention provides a general method by which the regions of peptides likely to contain T-cell epitopes can be described. Modifications to the peptide in these regions have the potential to modify the MHC Class II binding characteristics.

[0128] According to another aspect of the present invention, T-cell epitopes can be predicted with greater accuracy by the use of a more sophisticated computational method which takes into account the interactions of peptides with models of MHC Class II alleles.

[0129] The computational prediction of T-cell epitopes present within a peptide according to this particular aspect contemplates the construction of models of at least 42 MHC Class II alleles based upon the structures of all known MHC Class II molecules and a method for the use of these models in the computational identification of T-cell epitopes, the construction of libraries of peptide backbones for each model in order to allow for the known variability in relative peptide backbone alpha carbon (Cα) positions, the construction of libraries of amino-acid side chain conformations for each backbone dock with each model for each of the 20 amino-acid alternatives at positions critical for the interaction between peptide and MHC Class II molecule, and the use of these libraries of backbones and side-chain conformations in conjunction with a scoring function to select the optimum backbone and side-chain conformation for a particular peptide docked

with a particular MHC Class II molecule and the derivation of a binding score from this interaction.

[0130] Models of MHC Class II molecules can be derived via homology modeling from a number of similar structures found in the Brookhaven Protein Data Bank ("PDB"). These may be made by the use of semi-automatic homology modeling software (Modeller, Sali A. & Blundell TL., 1993. *J. Mol Biol* 234:779-815) which incorporates a simulated annealing function, in conjunction with the CHARMm force-field for energy minimization (available from Molecular Simulations Inc., San Diego, Ca.). Alternative modeling methods can be utilized as well.

[0131] The present method differs significantly from other computational methods which use libraries of experimentally derived binding data of each amino-acid alternative at each position in the binding groove for a small set of MHC Class II molecules (Marshall, K.W., *et al., Biomed. Pept. Proteins Nucleic Acids*, 1(3):157-162) (1995) or yet other computational methods which use similar experimental binding data in order to define the binding characteristics of particular types of binding pockets within the groove, again using a relatively small subset of MHC Class II molecules, and then 'mixing and matching' pocket types from this pocket library to artificially create further 'virtual' MHC Class II molecules (Sturniolo T., et al., *Nat. Biotech,* 17(6): 555-561 (1999). Both prior methods suffer the major disadvantage that, due to the complexity of the assays and the need to synthesize large numbers of peptide variants, only a small number of MHC Class II molecules can be experimentally scanned. Therefore the first prior method can only make predictions for a small number of MHC Class II molecules. The second prior method also makes the assumption that a pocket lined with similar amino-acids in one molecule will have the same binding characteristics when in the context of a different Class II allele and suffers further disadvantages in that only those MHC Class II molecules can be 'virtually' created which contain pockets contained within the pocket library. Using the modeling approach described herein, the structure of any number and type of MHC Class II molecules can be deduced, therefore alleles can be specifically selected to be representative of the global population. In addition, the number of MHC

Class II molecules scanned can be increased by making further models further than
having to generate additional data via complex experimentation.

[0132] The use of a backbone library allows for variation in the positions of the Cα
atoms of the various peptides being scanned when docked with particular MHC Class II
molecules. This is again in contrast to the alternative prior computational methods
described above which rely on the use of simplified peptide backbones for scanning
amino-acid binding in particular pockets. These simplified backbones are not likely to be
representative of backbone conformations found in 'real' peptides leading to inaccuracies
in prediction of peptide binding. The present backbone library is created by superposing
the backbones of all peptides bound to MHC Class II molecules found within the Protein
Data Bank and noting the root mean square (RMS) deviation between the Cα atoms of
each of the eleven amino-acids located within the binding groove. While this library can
be derived from a small number of suitable available mouse and human structures
(currently 13), in order to allow for the possibility of even greater variability, the RMS
figure for each C"-α position is increased by 50%. The average Cα position of each
amino-acid is then determined and a sphere drawn around this point whose radius equals
the RMS deviation at that position plus 50%. This sphere represents all allowed Cα
positions.

[0133] Working from the Cα with the least RMS deviation (that of the amino-acid in
Pocket 1 as mentioned above, equivalent to Position 2 of the 11 residues in the binding
groove), the sphere is three-dimensionally gridded, and each vertex within the grid is then
used as a possible location for a Cα of that amino-acid. The subsequent amide plane,
corresponding to the peptide bond to the subsequent amino-acid is grafted onto each of
these Cαs and the φ and ψ angles are rotated step-wise at set intervals in order to position
the subsequent Cα. If the subsequent Cα falls within the 'sphere of allowed positions'
for this Cα than the orientation of the dipeptide is accepted, whereas if it falls outside the
sphere then the dipeptide is rejected. This process is then repeated for each of the
subsequent Cα positions, such that the peptide grows from the Pocket 1 Cα 'seed', until
all nine subsequent Cαs have been positioned from all possible permutations of the

preceding Cαs. The process is then repeated once more for the single Cα preceding pocket 1 to create a library of backbone Cα positions located within the binding groove.

[0134] The number of backbones generated is dependent upon several factors: The size of the 'spheres of allowed positions'; the fineness of the gridding of the 'primary sphere' at the Pocket 1 position; the fineness of the step-wise rotation of the ϕ and ψ angles used to position subsequent Cαs. Using this process, a large library of backbones can be created. The larger the backbone library, the more likely it will be that the optimum fit will be found for a particular peptide within the binding groove of an MHC Class II molecule. Inasmuch as all backbones will not be suitable for docking with all the models of MHC Class II molecules due to clashes with amino-acids of the binding domains, for each allele a subset of the library is created comprising backbones which can be accommodated by that allele. The use of the backbone library, in conjunction with the models of MHC Class II molecules creates an exhaustive database consisting of allowed side chain conformations for each amino-acid in each position of the binding groove for each MHC Class II molecule docked with each allowed backbone. This data set is generated using a simple steric overlap function where a MHC Class II molecule is docked with a backbone and an amino-acid side chain is grafted onto the backbone at the desired position. Each of the rotatable bonds of the side chain is rotated step-wise at set intervals and the resultant positions of the atoms dependent upon that bond noted. The interaction of the atom with atoms of side-chains of the binding groove is noted and positions are either accepted or rejected according to the following criteria: The sum total of the overlap of all atoms so far positioned must not exceed a pre-determined value. Thus the stringency of the conformational search is a function of the interval used in the step-wise rotation of the bond and the pre-determined limit for the total overlap. This latter value can be small if it is known that a particular pocket is rigid, however the stringency can be relaxed if the positions of pocket side-chains are known to be relatively flexible. Thus allowances can be made to imitate variations in flexibility within pockets of the binding groove. This conformational search is then repeated for every amino-acid at every position of each backbone when docked with each of the MHC Class II molecules to create the exhaustive database of side-chain conformations.

[0135]   A suitable mathematical expression is used to estimate the energy of binding between models of MHC Class II molecules in conjunction with peptide ligand conformations which have to be empirically derived by scanning the large database of backbone/side-chain conformations described above. Thus a protein is scanned for potential T-cell epitopes by subjecting each possible peptide of length varying between 9 and 20 amino-acids (although the length is kept constant for each scan) to the following computations: An MHC Class II molecule is selected together with a peptide backbone allowed for that molecule and the side-chains corresponding to the desired peptide sequence are grafted on. Atom identity and interatomic distance data relating to a particular side-chain at a particular position on the backbone are collected for each allowed conformation of that amino-acid (obtained from the database described above). This is repeated for each side-chain along the backbone and peptide scores derived using a scoring function. The best score for that backbone is retained and the process repeated for each allowed backbone for the selected model. The scores from all allowed backbones are compared and the highest score is deemed to be the peptide score for the desired peptide in that MHC Class II model. This process is then repeated for each model with every possible peptide derived from the protein being scanned, and the scores for peptides versus models are displayed.

[0136]   In the context of the present invention, each ligand presented for the binding affinity calculation is an amino-acid segment selected from a peptide or protein as discussed above. Thus, the ligand is a selected stretch of amino acids about 9 to 20 amino acids in length derived from a peptide, polypeptide or protein of known sequence. The terms "amino acids" and "residues" are hereinafter regarded as equivalent terms. The ligand, in the form of the consecutive amino acids of the peptide to be examined grafted onto a backbone from the backbone library, is positioned in the binding cleft of an MHC Class II molecule from the MHC Class II molecule model library via the coordinates of the C"-α atoms of the peptide backbone and an allowed conformation for each side-chain is selected from the database of allowed conformations. The relevant atom identities and interatomic distances are also retrieved from this database and used to calculate the peptide binding score. Ligands with a high binding affinity for the MHC Class II binding

pocket are flagged as candidates for site-directed mutagenesis. Amino-acid substitutions are made in the flagged ligand (and hence in the protein of interest) which is then retested using the scoring function in order to determine changes which reduce the binding affinity below a predetermined threshold value. These changes can then be incorporated into the protein of interest to remove T-cell epitopes.

[0137] Binding between the peptide ligand and the binding groove of MHC Class II molecules involves non-covalent interactions including, but not limited to: hydrogen bonds, electrostatic interactions, hydrophobic (lipophilic) interactions and Van der Waal's interactions. These are included in the peptide scoring function as described in detail below. It should be understood that a hydrogen bond is a non-covalent bond which can be formed between polar or charged groups and consists of a hydrogen atom shared by two other atoms. The hydrogen of the hydrogen donor has a positive charge where the hydrogen acceptor has a partial negative charge. For the purposes of peptide/protein interactions, hydrogen bond donors may be either nitrogens with hydrogen attached or hydrogens attached to oxygen or nitrogen. Hydrogen bond acceptor atoms may be oxygens not attached to hydrogen, nitrogens with no hydrogens attached and one or two connections, or sulphurs with only one connection. Certain atoms, such as oxygens attached to hydrogens or imine nitrogens (e.g. C=NH) may be both hydrogen acceptors or donors. Hydrogen bond energies range from 3 to 7 Kcal/mol and are much stronger than Van der Waal's bonds, but weaker than covalent bonds. Hydrogen bonds are also highly directional and are at their strongest when the donor atom, hydrogen atom and acceptor atom are co-linear. Electrostatic bonds are formed between oppositely charged ion pairs and the strength of the interaction is inversely proportional to the square of the distance between the atoms according to Coulomb's law. The optimal distance between ion pairs is about 2.8Å. In protein/peptide interactions, electrostatic bonds may be formed between arginine, histidine or lysine and aspartate or glutamate. The strength of the bond will depend upon the pKa of the ionizing group and the dielectric constant of the medium although they are approximately similar in strength to hydrogen bonds.

[0138]   Lipophilic interactions are favorable hydrophobic-hydrophobic contacts that occur between he protein and peptide ligand. Usually, these will occur between hydrophobic amino acid side chains of the peptide buried within the pockets of the binding groove such that they are not exposed to solvent. Exposure of the hydrophobic residues to solvent is highly unfavorable since the surrounding solvent molecules are forced to hydrogen bond with each other forming cage-like clathrate structures. The resultant decrease in entropy is highly unfavorable. Lipophilic atoms may be sulphurs which are neither polar nor hydrogen acceptors and carbon atoms which are not polar.

[0139]   Van der Waal's bonds are non-specific forces found between atoms which are 3-4Å apart. They are weaker and less specific than hydrogen and electrostatic bonds. The distribution of electronic charge around an atom changes with time and, at any instant, the charge distribution is not symmetric. This transient asymmetry in electronic charge induces a similar asymmetry in neighboring atoms. The resultant attractive forces between atoms reaches a maximum at the Van der Waal's contact distance but diminishes very rapidly at about 1Å to about 2Å. Conversely, as atoms become separated by less than the contact distance, increasingly strong repulsive forces become dominant as the outer electron clouds of the atoms overlap. Although the attractive forces are relatively weak compared to electrostatic and hydrogen bonds (about 0.6 Kcal/mol), the repulsive forces in particular may be very important in determining whether a peptide ligand may bind successfully to a protein.

[0140]   In one embodiment, the Böhm scoring function (SCORE1 approach) is used to estimate the binding constant. (Böhm, H.J., *J. Comput Aided Mol. Des.*, 8(3):243-256 (1994) which is hereby incorporated in its entirety). In another embodiment, the scoring function (SCORE2 approach) is used to estimate the binding affinities as an indicator of a ligand containing a T-cell epitope (Böhm, H.J., *J. Comput Aided Mol. Des.*, 12(4):309-323 (1998) which is hereby incorporated in its entirety). However, the Böhm scoring functions as described in the above references are used to estimate the binding affinity of a ligand to a protein where it is already known that the ligand successfully binds to the protein and the protein/ligand complex has had its structure solved, the solved structure being present in the Protein Data Bank ("PDB"). Therefore, the scoring function has

been developed with the benefit of known positive binding data. In order to allow for discrimination between positive and negative binders, a repulsion term must be added to the equation. In addition, a more satisfactory estimate of binding energy is achieved by computing the lipophilic interactions in a pairwise manner rather than using the area based energy term of the above Böhm functions. Therefore, in a preferred embodiment, the binding energy is estimated using a modified Böhm scoring function. In the modified Böhm scoring function, the binding energy between protein and ligand (ΔG_{bind}) is estimated considering the following parameters: The reduction of binding energy due to the overall loss of translational and rotational entropy of the ligand (ΔG_0); contributions from ideal hydrogen bonds (ΔG_{hb}) where at least one partner is neutral; contributions from unperturbed ionic interactions (ΔG_{ionic}); lipophilic interactions between lipophilic ligand atoms and lipophilic acceptor atoms (ΔG_{lipo}); the loss of binding energy due to the freezing of internal degrees of freedom in the ligand, i.e., the freedom of rotation about each C-C bond is reduced (ΔG_{rot}); the energy of the interaction between the protein and ligand (E_{vdW}). Consideration of these terms gives <u>equation 1</u>:

$$(\Delta G_{bind}) = (\Delta G_0) + (\Delta G_{hb} \times N_{hb}) + (\Delta G_{ionic} \times N_{ionic}) + (\Delta G_{lipo} \times N_{lipo}) + (\Delta G_{rot} + N_{rot}) + (E_{vdW}).$$

Where N is the number of qualifying interactions for a specific term and, in one embodiment, ΔG_0, ΔG_{hb}, ΔG_{ionic}, ΔG_{lipo} and ΔG_{rot} are constants which are given the values: 5.4, -4.7, -4.7, -0.17, and 1.4, respectively.

The term N_{hb} is calculated according to <u>equation 2</u>:

$$N_{hb} = \Sigma_{h\text{-bonds}} f(\Delta R, \Delta\alpha) \times f(N_{neighb}) \times f_{pcs}$$

$f(\Delta R, \Delta\alpha)$ is a penalty function which accounts for large deviations of hydrogen bonds from ideality and is calculated according to <u>equation 3</u>:

$$f(\Delta R, \Delta\text{-}\alpha) = f1(\Delta R) \times f2(\Delta\alpha)$$

Where: $f1(\Delta R)$ $= 1$ if $\Delta R <= TOL$

        or $= 1 - (\Delta R - TOL)/0.4$ if $\Delta R <= 0.4 + TOL$

        or $= 0$ if $\Delta R > 0.4 + TOL$

And: $f2(\Delta\alpha)$ $= 1$ if $\Delta\alpha < 30°$

or $= 1-(\Delta\alpha - 30)/50$ if $\Delta\alpha <= 80°$

or $= 0$ if $\Delta\alpha > 80°$

TOL is the tolerated deviation in hydrogen bond length $= 0.25\text{Å}$

ΔR is the deviation of the H-O/N hydrogen bond length from the ideal value $= 1.9\text{Å}$

$\Delta\alpha$ is the deviation of the hydrogen bond angle $\angle$ N/O-H..O/N from its idealized value of 180°

$f(N_{neighb})$ distinguishes between concave and convex parts of a protein surface and therefore assigns greater weight to polar interactions found in pockets rather than those found at the protein surface. This function is calculated according to equation 4 below:

$f(N_{neighb}) = (N_{neighb}/N_{neighb,0})^{\alpha}$ where $\alpha = 0.5$

N_{neighb} is the number of non-hydrogen protein atoms that are closer than 5Å to any given protein atom.

$N_{neighb,0}$ is a constant $= 25$

f_{pcs} is a function which allows for the polar contact surface area per hydrogen bond and therefore distinguishes between strong and weak hydrogen bonds and its value is determined according to the following criteria:

$f_{pcs} = \beta$ when $A_{polar}/N_{HB} < 10 \text{ Å}^2$

or $f_{pcs} = 1$ when $A_{polar}/N_{HB} > 10 \text{ Å}^2$

A_{polar} is the size of the polar protein-ligand contact surface

N_{HB} is the number of hydrogen bonds

β is a constant whose value $= 1.2$

For the implementation of the modified Böhm scoring function, the contributions from ionic interactions, ΔG_{ionic}, are computed in a similar fashion to those from hydrogen bonds described above since the same geometry dependency is assumed.

The term N_{lipo} is calculated according to equation 5 below:

N_{lipo} $= \Sigma_{lL} f(r_{lL})$

$f(r_{lL})$ is calculated for all lipophilic ligand atoms, l, and all lipophilic protein atoms, L, according to the following criteria:

39

$f(r_{1L})$ =1 when r_{1L} <= R1$f(r_{1L})$ =(r_{1L} - R1)/(R2-R1) when R2 <r_{1L} > R1

$f(r_{1L})$ =0 when r_{1L} >= R2

Where: R1 = r_1^{vdw} + r_L^{vdw} + 0.5

and R2 = R1 + 3.0

and r_1^{vdw} is the Van der Waal's radius of atom 1

and r_L^{vdw} is the Van der Waal's radius of atom L

The term N_{rot} is the number of rotable bonds of the amino acid side chain and is taken to be the number of acyclic sp^3 - sp^3 and sp^3 - sp^2 bonds. Rotations of terminal -CH_3 or -NH_3 are not taken into account.

The final term, E_{vdw}, is calculated according to equation 6 below:

$$E_{vdw} = \varepsilon_1\varepsilon_2 ((r_1^{vdw} +r_2^{vdw})^{12}/r^{12} - (r_1^{vdw} +r_2^{vdw})^6/r^6), \text{ where:}$$

ε_1 and ε_2 are constants dependent upon atom identity

$r_1^{vdw} +r_2^{vdw}$ are the Van der Waal's atomic radii

r is the distance between a pair of atoms.

[0141] With regard to Equation 6, in one embodiment, the constants ε_1 and ε_2 are given the atom values: C: 0.245, N: 0.283, O: 0.316, S: 0.316, respectively (i.e. for atoms of Carbon, Nitrogen, Oxygen and Sulphur, respectively). With regards to equations 5 and 6, the Van der Waal's radii are given the atom values C: 1.85, N: 1.75, O: 1.60, S: 2.00Å.

[0142] It should be understood that all predetermined values and constants given in the equations above are determined within the constraints of current understandings of protein ligand interactions with particular regard to the type of computation being undertaken herein. Therefore, it is possible that, as this scoring function is refined further, these values and constants may change hence any suitable numerical value which gives the desired results in terms of estimating the binding energy of a protein to a ligand may be used and hence fall within the scope of the present invention.

[0143] As described above, the scoring function is applied to data extracted from the database of side-chain conformations, atom identities, and interatomic distances. For the purposes of the present description, the number of MHC Class II molecules included in this database is 42 models plus four solved structures. It should be apparent from the

above descriptions that the modular nature of the construction of the computational method of the present invention means that new models can simply be added and scanned with the peptide backbone library and side-chain conformational search function to create additional data sets which can be processed by the peptide scoring function as described above. This allows for the repertoire of scanned MHC Class II molecules to easily be increased, or structures and associated data to be replaced if data are available to create more accurate models of the existing alleles.

[0144] The present prediction method can be calibrated against a data set comprising a large number of peptides whose affinity for various MHC Class II molecules has previously been experimentally determined. By comparison of calculated versus experimental data, a cut of value can be determined above which it is known that all experimentally determined T-cell epitopes are correctly predicted.

[0145] It should be understood that, although the above scoring function is relatively simple compared to some sophisticated methodologies that are available, the calculations are performed extremely rapidly. It should also be understood that the objective is not to calculate the true binding energy *per se* for each peptide docked in the binding groove of a selected MHC Class II protein. The underlying objective is to obtain comparative binding energy data as an aid to predicting the location of T-cell epitopes based on the primary structure (i.e. amino acid sequence) of a selected protein. A relatively high binding energy or a binding energy above a selected threshold value would suggest the presence of a T-cell epitope in the ligand. The ligand may then be subjected to at least one round of amino-acid substitution and the binding energy recalculated. Due to the rapid nature of the calculations, these manipulations of the peptide sequence can be performed interactively within the program's user interface on cost-effectively available computer hardware. Major investment in computer hardware is thus not required.

[0146] It would be apparent to one skilled in the art that other available software could be used for the same purposes. In particular, more sophisticated software which is capable of docking ligands into protein binding-sites may be used in conjunction with energy minimization. Examples of docking software are: DOCK (Kuntz *et al.*, *J. Mol. Biol.*, 161:269-288 (1982)), LUDI (Böhm, H.J., *J. Comput Aided Mol. Des.*, 8:623-632 (1994))

and FLEXX (Rarey M., et al., ISMB, 3:300-308 (1995)). Examples of molecular modeling and manipulation software include: AMBER (Tripos) and CHARMm (Molecular Simulations Inc.). The use of these computational methods would severely limit the throughput of the method of this invention due to the lengths of processing time required to make the necessary calculations. However, it is feasible that such methods could be used as a 'secondary screen' to obtain more accurate calculations of binding energy for peptides which are found to be 'positive binders' via the method of the present invention. The limitation of processing time for sophisticated molecular mechanic or molecular dynamic calculations is one which is defined both by the design of the software which makes these calculations and the current technology limitations of computer hardware. It may be anticipated that, in the future, with the writing of more efficient code and the continuing increases in speed of computer processors, it may become feasible to make such calculations within a more manageable time-frame. Further information on energy functions applied to macromolecules and consideration of the various interactions that take place within a folded protein structure can be found in: Brooks, B.R., *et al.*, *J. Comput. Chem.*, 4:187-217 (1983) and further information concerning general protein-ligand interactions can be found in: Dauber-Osguthorpe et al., *Proteins*4(1):31-47(1988), which are incorporated herein by reference in their entirety. Useful background information can also be found, for example, in Fasman, G.D., ed., *Prediction of Protein Structure and the Principles of Protein Conformation,* Plenum Press, New York, ISBN: 0-306 4313-9.

Equivalents

[0147]   The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.

Incorporation by Reference

[0148] All patents, patent applications, and scientific publications mentioned herein above are incorporated by reference into this application in their entirety.

CLAIMS

What is claimed is:

1 1.   A method for reducing the immunogenicity of a fusion protein, the method

2   comprising:

3   i.   identifying a candidate T-cell epitope within a junction region spanning a

4   fusion junction of a fusion protein; and,

5   ii. changing an amino acid within the junction region to reduce the ability of the

6   candidate T-cell epitope to interact with a T cell receptor.

1 2.   A fusion protein produced by the method of claim 1.

1 3.   A method for reducing the immunogenicity of a fusion protein, the method

2   comprising changing a candidate T-cell epitope within a junction region spanning

3   a fusion junction of a fusion protein to reduce the ability of the candidate T-cell

4   epitope to interact with a T cell receptor:

5   i.   T-cell epitope

1 4.   A fusion protein produced by the method of claim 3.

1 5.   A method for reducing the immunogenicity of a fusion protein, the method

2   comprising introducing a glycosylation site within the junction region spanning

3   the

4   fusion junction.

1 6.   A method for reducing the immunogenicity of a fusion protein, the method

2   comprising introducing a glycosylation site within 10 amino acids of the fusion

3   junction.

1 7.   A method of claim 5, the method comprising introducing a glycosylation site

2   within 5 amino acids of the fusion junction.

1 8.   A method of claim 5, the method comprising introducing a glycosylation site

2   within 2 amino acids of the fusion junction.

1 9.   A method for reducing the immunogenicity of a fusion protein, the method

2   comprising the steps of:

3   i.   providing a fusion protein with a junction region comprising a substituted

4   amino acid; and

5   ii.   assaying said fusion protein in an immunogenicity assay.

1   10.   A fusion protein produced by the method of claim 5, 6, 7, 8, or 9.

1   11.   A method of claim 5-8, wherein the glycosylation is an N-linked glycosylation.

1   12.   A method of claim 5-8, wherein the glycosylation is an O-linked glycosylation.

1   13.   A fusion protein of claim 2, 4, or 9, wherein the protein comprises an Ig

2   region.

1   14.   A fusion protein of claim 2, 4, or 9, wherein the protein comprises a serum

2   albumin region.

1   15.   A fusion protein of claim 2, 4, or 9, wherein the protein comprises a cytokine

2   activity.

1   16.   A fusion protein of claim 2, 4, or 9, wherein the protein comprises a hormone

2   activity.

1   17.   A fusion protein of claim 13, wherein the Ig region, the Ig region comprises

2   sequences of more than one antibody isotype.

1   18.   A fusion protein with reduced immunogenicity comprising

2   a first protein; and,

3   a second protein linked to said first protein via a fusion junction,

4   wherein the amino acid sequence of a junction region surrounding the fusion

5 junction is modified to remove a non-self T-cell epitope.

1   19.   The fusion protein of claim 18, wherein the junction region comprises between

2   1 and 25 amino acids.

1   20.   The fusion protein of claim 18, wherein the junction region comprises between

2   1 and 15 amino acids.

1   21.   The fusion protein of claim 18, wherein the junction region comprises between

2   1 and 9 amino acids.

1   22.   The fusion protein of claim 18, wherein the junction region comprises an N-

2   linked or an O-linked glycosylation site.

1   23.   The fusion protein of claim 18, wherein the junction region comprises a spacer
2   or linker.

1   24.   The fusion protein of claim 18, wherein the junction region comprises an Asn-
2   X-Ser/Thr-Gly-amino acid sequence, wherein X is any amino acid.

1   25.   The fusion protein of claim 18, wherein the first protein comprises an Ig
2   molecule or a fragment thereof.

1   26.   The fusion protein of claim 25, wherein the C-terminus of said Ig molecule or
2   fragment thereof is linked to the N-terminus of said second protein.

1   27.   The fusion protein of claim 18, wherein the junction region comprises an IgG
2   sequence having an ATAT amino acid sequence instead of an LSLS amino acid
3   sequence.

1   28.   The fusion protein of claim 25, wherein the Ig molecule or fragment thereof
2   comprises an Fc molecule.

3   29.   The fusion protein of claim 25, wherein the Ig molecule or fragment thereof

4   comprises amino acid sequences from two antibody isotypes.

1   30.   The fusion protein of claim 29, wherein the Ig molecule or fragment thereof

2   comprises IgG1 and IgG2 amino acid sequences.

1   31.   The fusion protein of claim 18, wherein the second protein has cytokine

2   activity.

1   32.   The fusion protein of claim 18, wherein the second protein has hormone

2   activity.

1   33.   The fusion protein of claim 18, wherein the first protein is an albumin protein.

1   34.   A method of reducing the immunogenicity of a fusion protein, the method

2   comprising:

3   i.   identifying an amino acid in a peptide in a junction region, wherein the

4   amino acid is selected from the group consisting of a leucine, a valine, an isoleucine, a

5   methionine, a phenylalanine, a tryptophan and a tyrosine; and

6   ii.   changing the amino acid in the peptide, such that the ability of the

7   peptide to bind to MHC Class II is reduced.

1   35.   A fusion protein produced by the method of claim 34.

1   36.   A nucleic acid encoding a fusion protein with reduced immunogenicity, the

2   fusion

3   protein comprising:

4   i.   a first protein;

5   ii.   a second protein linked to the first protein via a fusion junction,

6   wherein the amino acid sequence of a junction region spanning the

7   fusion junction is modified to remove a non-self T-cell epitope.

1   iii.

2   37.   A nucleic acid of claim 36, wherein the junction region comprises a linker.

SEQUENCE LISTING

<110> Lexigen Pharmaceuticals Corp.

<120> Reducing the Immunogenicity of Fusion Proteins

<130> LEX-017PCT

<150> US 60/280,625
<151> 2001-03-30

<160> 59

<170> PatentIn version 3.0

<210> 1
<211> 330
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> human Ig gamma heavy chain C region

<400> 1

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1   5   10   15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
            20   25   30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
            35   40   45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
        50   55   60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65   70   75   80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
                85   90   95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys
            100   105   110

Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
            115   120   125

Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
        130   135   140

Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp
145   150   155   160

Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
                165   170   175

Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu

                    180   185   190

His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
        195   200   205

Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly
        210   215   220

Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu
225   230   235   240

Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
                245   250   255

Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn
                260   265   270

Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
            275   280   285

Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
    290   295   300

Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
305   310   315   320

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
                325   330

<210>   2
<211>   326
<212>   PRT
<213>   Homo sapiens

<220>
<221>   misc_feature
<223>   human Ig gamma-2 chain C region

<400>   2

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1   5   10   15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
            20   25   30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
        35   40   45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
    50   55   60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Asn Phe Gly Thr Gln Thr
65   70   75   80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
                85   90   95

Thr Val Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro
                100   105   110

Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
        115   120   125

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
    130   135   140

Val Ser His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly
145   150   155   160

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn
                165   170   175

Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Val His Gln Asp Trp
            180   185   190

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro
        195   200   205

Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro Arg Glu
    210   215   220

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
225   230   235   240

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
            245   250   255

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
            260   265   270

Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
        275   280   285

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
    290   295   300

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
305   310   315   320

Ser Leu Ser Pro Gly Lys
                325

<210>   3
<211>   362
<212>   PRT
<213>   Homo sapiens

<220>
<221>   misc_feature
<223>   human Ig3 constant region

<400>   3

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1   5   10   15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
            20   25   30

```
Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
        35                  40                  45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
    50                  55                  60                    .

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65                  70                  75                      80

Tyr Thr Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
            85                  90                  95

Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro
            100                 105                 110

Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
        115                 120                 125

Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys
    130                 135                 140

Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
145                 150                 155                 160

Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
            165                 170                 175

Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Lys Trp
            180                 185                 190

Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Trp Glu
            195                 200                 205

Glu Gln Tyr Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Leu
    210                 215                 220

His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
225                 230                 235                 240

Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly
            245                 250                 255

Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu
            260                 265                 270

Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
        275                 280                 285

Pro Ser Asp Ile Ala Met Glu Trp Glu Ser Ser Gly Gln Pro Glu Asn
    290                 295                 300

Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
305                 310                 315                 320

Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
            325                 330                 335

Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
            340                 345                 350

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
```

```
        355                     360
```

<210> 4
<211> 327
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Ig gamma-4 chain C region

<400> 4

```
Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1               5                   10                  15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
            20                  25                  30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
            35                  40                  45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
    50                  55                  60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys Thr
65                  70                  75                  80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
                85                  90                  95

Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro
            100                 105                 110

Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
            115                 120                 125

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
    130                 135                 140

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
145                 150                 155                 160

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
                165                 170                 175

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
            180                 185                 190

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
    195                 200                 205

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
    210                 215                 220

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys
225                 230                 235                 240

Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
                245                 250                 255
```

```
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
             260             265             270

Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
         275             280             285

Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
     290             295             300

Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
305             310             315             320

Leu Ser Leu Ser Leu Gly Lys
             325
```

```
<210>   5
<211>   9
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope

<400>   5

Lys Ser Leu Ser Leu Ser Pro Gly Lys
1               5

<210>   6
<211>   9
<212>   PRT
<213>   Artificial sequence

<220>
<223>   mutated potential T cell epitope

<400>   6

Lys Ser Ala Thr Ala Thr Pro Gly Lys
1               5

<210>   7
<211>   9
<212>   PRT
<213>   Artificial sequence

<220>
<223>   mutated potential T cell epitope

<400>   7

Lys Ser Ala Thr Ala Thr Pro Gly Ala
1               5

<210>   8
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
```

<223>   sequence near the C-terminus of CH3 domain of the IgA Fc region

<400>   8

Gln Lys Thr Ile Asp Arg Leu Ala Gly Lys Pro Thr His
1   5   10

<210>   9
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   mutated sequence near the C-terminus of CH3 domain of the IgA Fc
         region

<400>   9

Gln Lys Thr Ala Asp Arg Thr Ala Gly Lys Pro Thr His
1   5   10

<210>   10
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   deimmunized sequence in an IgA-X fusion

<400>   10

Gln Lys Thr Pro Thr Arg Thr Ala Gly Lys Pro Thr His
1   5   10

<210>   11
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   deimmunized sequence in an IgA-X fusion

<400>   11

Gln Lys Thr Pro Thr Arg Pro Ala Gly Lys Pro Thr His
1   5   10

<210>   12
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   deimmunized sequence in an IgA-X fusion

<400>   12

Gln Lys Thr Ala Thr Arg Pro Ala Gly Lys Pro Thr His
1   5   10

<210>   13

```
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in the HSA-IFNalpha junction

<400>   13

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
1               5                   10

<210>   14
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified C-terminus of albumin

<400>   14

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Thr Thr Ala
1               5                   10

<210>   15
<211>   30
<212>   PRT
<213>   Artificial sequence

<220>
<223>   sequence in an Fc fusion protein

<400>   15

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly
1               5                   10                  15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20                  25                  30

<210>   16
<211>   30
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified sequence in an Fc fusion protein

<400>   16

His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Lys Gly
1               5                   10                  15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20                  25                  30

<210>   17
<211>   9
<212>   PRT
<213>   Artificial sequence
```

```
<220>
<223>  junction sequence

<400>  17

Leu Ser Leu Ser Pro Gly Lys Ala Pro
1               5

<210>  18
<211>  9
<212>  PRT
<213>  Artificial sequence

<220>
<223>  modified junction sequence

<400>  18

Ala Thr Ala Thr Pro Gly Ala Ala Pro
1               5

<210>  19
<211>  9
<212>  PRT
<213>  Artificial sequence

<220>
<223>  modified junction sequence

<400>  19

Leu Asn Leu Ser Pro Gly Ala Ala Pro
1               5

<210>  20
<211>  13
<212>  PRT
<213>  Artificial sequence

<220>
<223>  synthetic peptide containing a reactive epitope

<400>  20

Lys Ser Leu Ser Leu Ser Pro Gly Lys Ala Pro Thr Ser
1               5                   10

<210>  21
<211>  13
<212>  PRT
<213>  Artificial sequence

<220>
<223>  modified synthetic peptide containing a reactive epitope

<400>  21

Lys Ser Ala Thr Ala Thr Pro Gly Lys Ala Pro Thr Ser
1               5                   10
```

```
<210>   22
<211>   39
<212>   PRT
<213>   Artificial sequence

<220>
<223>   albumin-CD4 junction sequence

<400>   22

Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala
1               5                   10                  15

Ala Leu Gly Leu Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val
            20                  25                  30

Glu Leu Thr Cys Thr Ala Ser
        35

<210>   23
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in HSA-IFNalpha fusion

<400>   23

Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Cys
1               5                   10

<210>   24
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in HSA-IFNalpha fusion

<400>   24

Leu Gly Leu Cys Asp Leu Pro Gln Thr His Ser Leu Gly
1               5                   10

<210>   25
<211>   5
<212>   PRT
<213>   Artificial sequence

<220>
<223>   C-terminus albumin sequence

<400>   25

Ala Ala Leu Gly Leu
1               5

<210>   26
<211>   5
<212>   PRT
```

```
<213>  Artificial sequence

<220>
<223>  mutated C-terminus albumin sequence

<400>  26

Thr Ala Thr Thr Ala
1               5

<210>  27
<211>  19
<212>  PRT
<213>  Artificial sequence

<220>
<223>  modified albumin junction region

<400>  27

Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Thr Ala
1               5                   10                  15

Thr Thr Ala


<210>  28
<211>  21
<212>  PRT
<213>  Artificial sequence

<220>
<223>  non-mutant Fc sequence

<400>  28

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1               5                   10                  15

Pro Glu Leu Leu Gly
            20

<210>  29
<211>  52
<212>  PRT
<213>  Artificial sequence

<220>
<223>  GLP-1-mutant Fc fusion junction

<400>  29

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1               5                   10                  15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
            20                  25                  30

Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
            35                  40                  45
```

Glu Leu Leu Gly
    50

<210>   30
<211>   41
<212>   PRT
<213>   Artificial sequence

<220>
<223>   GLP-1-normal Fc fusion junction

<400>   30

Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
1   5   10   15

Lys Gly Arg Gly Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20   25   30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35   40

<210>   31
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope at the GLP-1-Fc fusion

<400>   31

Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
1   5   10

<210>   32
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope at the GLP-1-Fc fusion junction

<400>   32

Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu Pro
1   5   10

<210>   33
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in GLP-1-Fc fusion junction

<400>   33

Ala Trp Leu Val Lys Gly Arg Gly Glu Pro Lys Ser Ser
1   5   10

```
<210>   34
<211>   52
<212>   PRT
<213>   Artificial sequence

<220>
<223>   deimmunized GLP-1Fc fusion junction

<400>   34

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1               5                   10                  15

Gln Ala Ala Lys Glu Phe Ala Ala Trp Ala Val Thr Gly Thr Gly Glu
            20                  25                  30

Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
        35                  40                  45

Glu Leu Leu Gly
    50

<210>   35
<211>   41
<212>   PRT
<213>   Artificial sequence

<220>
<223>   GLP-1-Fc fusion junction with a glycosylation site

<400>   35

Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
1               5                   10                  15

Lys Gly Arg Asn Gly Ser Lys Ser Ser Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35                  40

<210>   36
<211>   41
<212>   PRT
<213>   Artificial sequence

<220>
<223>   TNF-R-gamma-1 fusion junction

<400>   36

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
1               5                   10                  15

Ser Thr Gly Asp Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35                  40

<210>   37
```

```
<211>    41
<212>    PRT
<213>    Artificial sequence

<220>
<223>    TNF-R-Fc fusion junction·

<400>    37

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
1                5                   10                  15

Ser Thr Gly Asn Gly Ser Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35                  40

<210>    38
<211>    40
<212>    PRT
<213>    Artificial sequence

<220>
<223>    Fc-IL12p35 fusion junction

<400>    38

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
1                5                   10                  15

Ser Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
            20                  25                  30

Phe Pro Cys Leu His His Ser Gln
        35                  40

<210>    39
<211>    40
<212>    PRT
<213>    Artificial sequence

<220>
<223>    IL-12p40-IL2 fusion junction

<400>    39

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1                5                   10                  15

Val Pro Cys Ser Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
            20                  25                  30

Gln Leu Glu His Leu Leu Leu Asp
        35                  40

<210>    40
<211>    40
<212>    PRT
<213>    Artificial sequence
```

<220>
<223> modified Fc-IL12p35 fusion junction

<400> 40

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1   5   10   15

Thr Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
            20   25   30

Phe Pro Cys Leu His His Ser Gln
        35   40

<210> 41
<211> 40
<212> PRT
<213> Artificial sequence

<220>
<223> modified IL12p40-IL2 fusion junction

<400> 41

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1   5   10   15

Val Pro Cys Ser Asn Gly Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
            20   25   30

Gln Leu Glu His Leu Leu Leu Asp
        35   40

<210> 42
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> potential T cell epitope in IL12p40-IL fusion

<400> 42

Ser Glu Trp Ala Ser Val Pro Cys Ser Asn Gly Thr Ser
1   5   10

<210> 43
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> potential T cell epitope in IL12p40-IL2 fusion junction

<400> 43

Ala Ser Val Pro Cys Ser Asn Gly Thr Ser Ser Ser Thr
1   5   10

<210> 44
<211> 41

```
<212>   PRT
<213>   Artificial sequence

<220>
<223>   IL4-Fc fusion junction

<400>   44

Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
1               5                   10                  15

Lys Cys Ser Ser Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35                  40

<210>   45
<211>   40
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Fc-GMCSF fusion junction

<400>   45

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
1               5                   10                  15

Ser Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
            20                  25                  30

Glu His Val Asn Ala Ile Gln Glu
        35                  40

<210>   46
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T-cell epitope at IL4-Fc fusion junction

<400>   46

Glu Lys Tyr Ser Lys Cys Ser Ser Glu Pro Lys Ser Cys
1               5                   10

<210>   47
<211>   41
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified IL4-Fc fusion

<400>   47

Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
1               5                   10                  15
```

```
Lys Cys Ser Ser Thr Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
        35                  40
```

```
<210>   48
<211>   40
<212>   PRT
<213>   Artificial sequence

<220>
<223>   deimmunized Fc-GMCSF fusion junction

<400>   48
```

```
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1               5                   10                  15

Thr Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
            20                  25                  30

Glu His Val Asn Ala Ile Gln Glu
        35                  40
```

```
<210>   49
<211>   35
<212>   PRT
<213>   Artificial sequence

<220>
<223>   IgG2CH1-IgG1hinge fusion junction

<400>   49
```

```
Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
1               5                   10                  15

Asp Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro
        35
```

```
<210>   50
<211>   35
<212>   PRT
<213>   Artificial sequence

<220>
<223>   IgG1hinge-IgG2CH2 fusion junction

<400>   50
```

```
Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1               5                   10                  15

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
            20                  25                  30
```

Asp Thr Leu
        35

<210>   51
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in the IgG2CH1-IgG1 hinge fusion junctio

<400>   51

Thr Lys Val Asp Lys Thr Val Glu Pro Lys Ser Cys Asp
1   5   10

<210>   52
<211>   13
<212>   PRT
<213>   Artificial sequence

<220>
<223>   potential T cell epitope in the IgG2CH1-IgG1 hinge fusion junctio

<400>   52

Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
1   5   10

<210>   53
<211>   35
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified IgG2CH1-IgG1hinge fusion junction

<400>   53

Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Ala
1   5   10   15

Asp Lys Thr Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
            20   25   30

Pro Cys Pro
        35

<210>   54
<211>   35
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified IgG2CH1-IgG1hinge fusion junction

<400>   54

Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val

```
1                    5                      10                     15

Asp Lys Thr Val Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro
            20                  25                  30

Pro Cys Pro
        35


<210>   55
<211>   35
<212>   PRT
<213>   Artificial sequence

<220>
<223>   modified IgG1hinge-IgG2CH2 fusion junction

<400>   55

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1               5                   10                  15

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
            20                  25                  30

Asp Thr Leu
        35


<210>   56
<211>   166
<212>   PRT
<213>   Artificial sequence

<220>
<223>   mutant EPO sequence

<400>   56

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1               5                   10                  15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu Gly
            20                  25                  30

Pro Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
        35                  40                  45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
    50                  55                  60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
65                  70                  75                  80

Leu Val Asn Ser Ser Gln Pro Cys Glu Gly Leu Gln Leu His Val Asp
                85                  90                  95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
            100                 105                 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
        115                 120                 125
```

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
    130   135   140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145   150   155   160

Cys Arg Thr Gly Asp Arg
                165

<210> 57
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> CH3-EPO fusion junction

<400> 57

Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Ala Ala Pro Pro Arg Leu
1   5   10   15

Ile

<210> 58
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> IgG2 CH3 sequence

<400> 58

Lys Ser Leu Ser Leu Ser Pro Gly
1   5

<210> 59
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> modified IgG2CH3 sequence

<400> 59

Lys Ser Ala Thr Ala Thr Pro Gly
1   5